



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 103865

To: Frank Prats
Location: CM1-11B01
Art Unit: 1651
Tuesday, September 23, 2003

Case Serial Number: 10/069908

From: Beverly Shears
Location: Biotech-Chem Library
CM1-1E05
Phone: 308-4994

beverly.shears@uspto.gov

Search Notes

CRFE

Access DB# 103865

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: FRANCISCO CHANDLER PRATS Examiner #: 7618 Date: 9-12-03
Art Unit: 1651 Phone Number 308-3665 Serial Number: 10/069 908
Mail Box and Bldg/Room Location: 11301 CM Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: MALTOGENIC AMYLASE - MODIFIED STARCH DERIVATIVE

Inventors (please provide full names): SVEN PEDERSEN, HANNE VANG

Earliest Priority Filing Date: 9-1-99

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

PLEASE SEARCH SEQ ID NO. 1. IT ENCODES AN ENZYME CALLED A "MALTOGENIC AMYLASE" FIRST DISCLOSED IN EP 120 693.

I NEED TO KNOW IF THAT ENZYME, OR A "VARIANT" HAS BEEN USED TO DIGEST STARCH DERIVATIVES, SUCH AS OCTENYL SUCCINATE DERIVATIVES OF STARCH.

THANKS !

EXR PRATS

[Signature]

STAFF USE ONLY

Searcher: Beverly 29999

Searcher Phone #: _____

Searcher Location: _____

Date Searcher Picked Up: _____

Date Completed: 09-22-03

Searcher Prep & Review Time: 3

Clerical Prep Time: _____

Online Time: 20

Type of Search

NA Sequence (#) _____

AA Sequence (#) _____

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Orbit _____

Dr.Link _____

Lexis/Nexis _____

Sequence Systems _____

WWW/Internet _____

Other (specify) CGN

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 17, 2003, 07:32:22 ; Search time 5270 seconds
(without alignments)
16767.513 Million cell updates/sec

Title: US-10-069-908-1
Perfect score: 2160
Sequence: 1 atgaaaagaacacgtttc.....ctgttaoatggcaaaactag 2160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :

- 1: gb.ba.*
- 2: gb.btg.*
- 3: gb.in.*
- 4: gb.ov.*
- 5: gb.pat.*
- 6: gb.ph.*
- 7: gb.pl.*
- 8: gb.pr.*
- 9: gb.ro.*
- 10: gb.sts.*
- 11: gb.sy.*
- 12: gb.un.*
- 13: gb.vi.*
- 14: gb.ba.*
- 15: em.fun.*
- 16: em.hum.*
- 17: em.in.*
- 18: em.mus.*
- 19: em.ov.*
- 20: em.pat.*
- 21: em.ph.*
- 22: em.ro.*
- 23: em.sts.*
- 24: em.un.*
- 25: em.vi.*
- 26: em.btg.*
- 27: em.hum.*
- 28: em.in.*
- 29: em.mus.*
- 30: em.ov.*
- 31: em.pat.*
- 32: em.ph.*
- 33: em.ro.*
- 34: em.sts.*
- 35: em.un.*
- 36: em.vi.*
- 37: em.btg.*
- 38: em.hum.*
- 39: em.in.*
- 40: em.mus.*
- 41: em.ov.*
- 42: em.pat.*
- 43: em.ph.*
- 44: em.ro.*
- 45: em.sts.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2160	100.0	2160	6	ARI36896	ARI36896 Sequence
2	2160	100.0	2160	6	AR255841	AR255841 Sequence
3	2089.6	96.7	2726	1	BACMYLAA	M36539 B.stearothe
4	576	26.7	576	6	ARI39106	ARI39106 Sequence
5	395.8	18.3	2615	1	BCCGTASEX	L25256 Bacillus Q
6	383.4	17.8	2383	1	BCCGTASE	X68326 B.circulans
7	374.2	17.3	2516	1	BLCGTAG	X15752 Bacillus li
8	363.4	16.8	2489	1	BSCGTDNA	X66106 Bacillus sp
9	353	16.3	2197	1	BC251CGT	X78145 B. circulan
10	348.8	16.1	2058	6	E01516	E01516 DNA encodin
11	348.8	16.1	2139	6	E00874	E00874 DNA encodin
12	347.2	16.1	2058	6	E09507	E09507 DNA encodin
13	347.2	16.1	2058	6	E12945	E12945 DNA encodin
14	345	16.0	2996	1	BACMYRSD	M3302 Bacillus sp
15	345	16.0	2996	1	BACRSDA	D90112 Bacillus,sp
16	345	16.0	2996	6	E02907	E02907 DNA encodin
17	333.4	15.4	3025	1	AF047363	AF047363 Paenibaci
18	331.8	15.4	2631	1	BMCGTM	X59045 B.macerans
19	330.2	15.3	2414	1	BSGCT5	X59042 B.stearothe
20	328.6	15.2	2414	1	BSGCT1	X59044 B.stearothe
21	328.6	15.2	2414	6	E05456	E05456 DNA Sequence
22	327	15.1	327	6	ARI39105	ARI39105 Sequence
23	326.2	15.1	2566	6	E02162	E02162 Cyclomaltde
24	326.2	15.1	2648	1	BACCGTD	M28053 Bacillus sp
25	324.6	15.0	2659	6	E02671	AF302787 Bacillus
26	324.6	15.0	2659	6	E02671	E02671 DNA encodin
27	321.8	14.9	2040	6	E01514	E01514 DNA encodin
28	321.8	14.9	2040	6	E09506	E09506 DNA encodin
29	321.8	14.9	2040	6	E12944	E12944 DNA encodin
30	321.8	14.9	2040	6	E13426	E13426 DNA encodin
31	321.8	14.9	2040	6	E00873	E00873 DNA encodin
32	321.8	14.9	2133	6	E00873	E00873 DNA encodin
33	321.6	14.9	3048	1	BACCGTA	M17366 Bacillus sp
34	318.2	14.7	2133	6	AR095553	AR095553 Sequence
35	315.6	14.6	12854	1	TTFULSA	M57692 Thermoanaer
36	310.4	14.4	2661	1	BSU83799	U83799 Bacillus st
37	309.4	14.3	3438	1	BACMYE	M16657 B.circulans
38	306.6	14.2	13352	1	AE013135	AE013135 Thermoana
39	306.4	14.2	4794	1	TSCGTAG	Z35484 Thermoanaer
40	305.6	14.1	2463	1	BACCGTC	M19880 Bacillus sp
41	305.6	14.1	2686	6	E02670	E02670 DNA encodin
42	304	14.1	2687	6	E02161	E02161 Cyclomaltde
43	303.8	14.1	2502	1	BACCDG	M12777 B.mecerans
44	296	13.7	2061	6	E08061	E08061 DNA encodin
45	296	13.7	2061	6	I23834	I23834 Sequence 9

ALIGNMENTS

RESULT 1
LOCUS ARI36896 2160 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6162628.
ACCESSION ARI36896
VERSION ARI36896.1 GI:14478146
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Cherry,J., Svendsen,A., Andersen,C., Beier,L. and Frandsen,T.Peter.
TITLE Maltogenic alpha-amylase variants
JOURNAL Patent: US 6162628-A 1 19-DEC-2000;
FEATURES Location/Qualifiers

```
source      1. .2160
/organism="unknown"
BASE COUNT  571 a 499 c 572 g 518 t
ORIGIN
Query Match      100.0%; Score 2160; DB 6; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAGAAACCGCTTCTTATTTGTGGGACATGCTCTCTCATCGCTCTTCTGTC 60
DB 1 ATGAAAAGAAACCGCTTCTTATTTGTGGGACATGCTCTCTCATCGCTCTTCTGTC 60
QY 61 AGCGGTTCTCTTCGCTACAAATCCAAACCGCGCTGAAGCAGCAGTTCCGAAAGCGTCAA 120
DB 61 AGCGGTTCTCTTCGCTACAAATCCAAACCGCGCTGAAGCAGCAGTTCCGAAAGCGTCAA 120
QY 121 GGGACGCTGATTTACAGATTAATCATTTGACCGGTTTTCATGCGGACACGACGAAACAC 180
DB 121 GGGACGCTGATTTACAGATTAATCATTTGACCGGTTTTCATGCGGACACGACGAAACAC 180
QY 181 AATCTCCCAAAGTTTATGAGCTTTTACGATCCGACCAATCGAAGTGGAAATGTATTGG 240
DB 181 AATCTCCCAAAGTTTATGAGCTTTTACGATCCGACCAATCGAAGTGGAAATGTATTGG 240
QY 241 GCGGGGATCTGAGGGGTTCTGCAAAAACCTTCTTATCTTAAACAGCTGGCGGTAACG 300
DB 241 GCGGGGATCTGAGGGGTTCTGCAAAAACCTTCTTATCTTAAACAGCTGGCGGTAACG 300
QY 301 ACAATCTGTTGTCCTCCGTTTTTGGACAACTCTGTATACACTGGCGGCAACGATAACAG 360
DB 301 ACAATCTGTTGTCCTCCGTTTTTGGACAACTCTGTATACACTGGCGGCAACGATAACAG 360
QY 361 GCGTATCACGGAATCTGACGCGCGGANTTTAAACAGATTGAGGAACATTTCCGGAATGG 420
DB 361 GCGTATCACGGAATCTGACGCGCGGANTTTAAACAGATTGAGGAACATTTCCGGAATGG 420
QY 421 ACCCAATTTGACAGCTTGTGTAATGATGCTCACCAAAAGGAATCAAGGTGATTTGCGAC 480
DB 421 ACCCAATTTGACAGCTTGTGTAATGATGCTCACCAAAAGGAATCAAGGTGATTTGCGAC 480
QY 481 TTGTGCGCAATCATTCGACTCCTTTTAAAGCAAAAGATTCACCTTTTCGGAAGCGGC 540
DB 481 TTGTGCGCAATCATTCGACTCCTTTTAAAGCAAAAGATTCACCTTTTCGGAAGCGGC 540
QY 541 GCGCTCTACAAATGAGCACTTATGCGCAATTTTGTATGACGCAAAAGGCTAC 600
DB 541 GCGCTCTACAAATGAGCACTTATGCGCAATTTTGTATGACGCAAAAGGCTAC 600
QY 601 TTCCACCAATATGGGACATCAGCAACTGGGACGACCGGTACGAGCGCAATGGAAAAAC 660
DB 601 TTCCACCAATATGGGACATCAGCAACTGGGACGACCGGTACGAGCGCAATGGAAAAAC 660
QY 661 TTCAGGATCAGCGGTTTCTCGCTTCGCAATTTGTGCGAGAAATGGCAGATGCT 720
DB 661 TTCAGGATCAGCGGTTTCTCGCTTCGCAATTTGTGCGAGAAATGGCAGATGCT 720
QY 721 CAATACCTGACGATGCGGTTTCAATTTGTGTAGCAATGAGAGCGATGTTTCGGATT 780
DB 721 CAATACCTGACGATGCGGTTTCAATTTGTGTAGCAATGAGAGCGATGTTTCGGATT 780
QY 781 GATCGGTGAGCAATTTTAAATTCGGGTTTCTCCAAATCGTTGGCGGATAAATCTGTACAA 840
DB 781 GATCGGTGAGCAATTTTAAATTCGGGTTTCTCCAAATCGTTGGCGGATAAATCTGTACAA 840
QY 841 AAGAAGACATTTTCTGTTGGGGAATGTTACGAGATGATCCCGGACAGCGCAATCAT 900
DB 841 AAGAAGACATTTTCTGTTGGGGAATGTTACGAGATGATCCCGGACAGCGCAATCAT 900
QY 901 CTGGAAGAGTCCGGTACGCCAACACACAGCGGTGCAATGTGTGATTTTGTCTCAAC 960
DB 901 CTGGAAGAGTCCGGTACGCCAACACACAGCGGTGCAATGTGTGATTTTGTCTCAAC 960
```

```
QY 961 ACGGTGATTCGAAATGTGTTCGGCACATTTACGCAAAACGATGTACGATCTTTAAACAATG 1020
DB 961 ACGGTGATTCGAAATGTGTTCGGCACATTTACGCAAAACGATGTACGATCTTTAAACAATG 1020
QY 1021 GTGAACCAAACGGGGAACGAGTACAAATACAAAGAAAATCTTAATCACATTTATCGATAAC 1080
DB 1021 GTGAACCAAACGGGGAACGAGTACAAATACAAAGAAAATCTTAATCACATTTATCGATAAC 1080
QY 1081 CATGATATGTCAAGATTCTTTTCGGTAAATTCGAAACAGGCGAATTTGACAGCGCGCTT 1140
DB 1081 CATGATATGTCAAGATTCTTTTCGGTAAATTCGAAACAGGCGAATTTGACAGCGCGCTT 1140
QY 1141 GCTTTTCAATCTCACTTCGCGGGGTACGCCCTCCATCTATTTATGGAACCGAAACAATG 1200
DB 1141 GCTTTTCAATCTCACTTCGCGGGGTACGCCCTCCATCTATTTATGGAACCGAAACAATG 1200
QY 1201 GCAGCGCAATGACCCGCTACACCGGGGGATGATCCGCGTTTGTATACGACCAACACACC 1260
DB 1201 GCAGCGCAATGACCCGCTACACCGGGGGATGATCCGCGTTTGTATACGACCAACACACC 1260
QY 1261 GCGTTTAAAGAGGTGTCAACTCTGCGGGGTTCGCAAGGAACAATCGCGCGATCCAGTAC 1320
DB 1261 GCGTTTAAAGAGGTGTCAACTCTGCGGGGTTCGCAAGGAACAATCGCGCGATCCAGTAC 1320
QY 1321 GGCACACACCGAGCTTGGATCAACATGATCTTTTACATTTTATGAAACGGAATTTTTC 1380
DB 1321 GGCACACACCGAGCTTGGATCAACATGATCTTTTACATTTTATGAAACGGAATTTTTC 1380
QY 1381 AACGATGCTGTGTTGTTGGCCATCAATCGAAACACGCAATCTCTTATTCGATTTCCGGT 1440
DB 1381 AACGATGCTGTGTTGTTGGCCATCAATCGAAACACGCAATCTCTTATTCGATTTCCGGT 1440
QY 1441 TTGCGAGACGGCTTCGCAATGCGAGCTATCGCGATTAATCTGTACAGGCTGTGGGGGG 1500
DB 1441 TTGCGAGACGGCTTCGCAATGCGAGCTATCGCGATTAATCTGTACAGGCTGTGGGGGG 1500
QY 1501 AACGGGATTTCCGTTTCCAAATGAGGTGCGCTTCAGCTTCGCTTCGCTTCGAGCGCGTG 1560
DB 1501 AACGGGATTTCCGTTTCCAAATGAGGTGCGCTTCAGCTTCGCTTCGCTTCGAGCGCGTG 1560
QY 1561 TCTGTTGGCGATGACACATCCGCTTCAGCGCGCAATCCGATCGGTTGCTTCCAAAT 1620
DB 1561 TCTGTTGGCGATGACACATCCGCTTCAGCGCGCAATCCGATCGGTTGCTTCCAAAT 1620
QY 1621 ATGGGATTTCCGGTAAATGCTACGATGACGAGGAAAGTTTTTGGACGACGACGAGGA 1680
DB 1621 ATGGGATTTCCGGTAAATGCTACGATGACGAGGAAAGTTTTTGGACGACGACGAGGA 1680
QY 1681 ACCGTGACATTTGCGGAGTGCACGCGACTGTGAAATCTCTGGACATCCCAATCGGATTGAA 1740
DB 1681 ACCGTGACATTTGCGGAGTGCACGCGACTGTGAAATCTCTGGACATCCCAATCGGATTGAA 1740
QY 1741 GTGTACGTTCCCAATGSCCGCGGCTGACCGATCTGAAATCTACCGCGGCTGGAGTT 1800
DB 1741 GTGTACGTTCCCAATGSCCGCGGCTGACCGATCTGAAATCTACCGCGGCTGGAGTT 1800
QY 1801 TCCAGCAATCTGTATCTTCAATATTTTGTAGTGGAAACGACGACATCGGTTGTGTTACT 1860
DB 1801 TCCAGCAATCTGTATCTTCAATATTTTGTAGTGGAAACGACGACATCGGTTGTGTTACT 1860
QY 1861 GTGAAGTGGCGCTTCGACCACTGGGGGATTAAGATTTTACCTGTGCGGGCAACATACCG 1920
DB 1861 GTGAAGTGGCGCTTCGACCACTGGGGGATTAAGATTTTACCTGTGCGGGCAACATACCG 1920
QY 1921 GAATTCGGGAATTTGAGCAGCGATACGAGCGGAGCGCTTAAACATTCGCAAGGCGCCCTG 1980
DB 1921 GAATTCGGGAATTTGAGCAGCGATACGAGCGGAGCGCTTAAACATTCGCAAGGCGCCCTG 1980
QY 1981 CTGCGCCCAATTTATCCGATTTGTTTTATGTTATTCAGCGTTCCAGCAGGAAGACGATT 2040
DB 1981 CTGCGCCCAATTTATCCGATTTGTTTTATGTTATTCAGCGTTCCAGCAGGAAGACGATT 2040
QY 2041 CAATTCAAGTCTTCTCATAAGCGTTCGGATGGAACGATTCATGGGAGATGGTTTCGAAC 2100
```

Db 2041 CAATTCAAGTTCTTCATCAAGCTGCGGATGGAAGCAATCAATGGGAGATGTTTCGAAC 2100
Qy 2101 CAGTGGCCACAACCTCCACAGGGTGCAACCGGTAACTTACTGTACGTGGCAAACTAG 2160
Db 2101 CAGTGGCCACAACCTCCACAGGGTGCAACCGGTAACTTACTGTACGTGGCAAACTAG 2160
RESULT 2
AR255841
LOCUS AR255841 2160 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6482622.
ACCESSION AR255841
VERSION AR255841.1 GI:27305028
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Cherry,J.R., Svendsen,A., Andersen,C.D., Beier,L., Frandsen,T.P.
and Schafer,T.
TITLE Amylolytic enzyme variants
JOURNAL Patent: US 6482622-A 1 19-NOV-2002;
FEATURES Location/Qualifiers
source 1..2160
/organism="unknown"
BASE COUNT 571 a 499 c 572 g 518 t
ORIGIN
Query Match 100.0%; Score 2160; DB 6; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAAGAAACGGTCTTCTTTATTTGGGACGTAGTCTCCCTCATCGGTCTCTGTTTC 60
Db 1 ATGAAAAGAAACGGTCTTCTTTATTTGGGACGTAGTCTCTCATCGGTCTCTGTTTC 60
Qy 61 AGCGGTCTCTTCGGTACAACTCAAAACCGCTGAAGCCAGTCAGTCCGCAACGGTCAAA 120
Db 61 AGCGGTCTCTTCGGTACAACTCAAAACCGCTGAAGCCAGTCAGTCCGCAACGGTCAAA 120
Qy 121 GGGAGCGTATTTACAGATTATCATTTAGCCGGTCTTTAGTGGGACACGACGAACAAC 180
Db 121 GGGAGCGTATTTACAGATTATCATTTAGCCGGTCTTTAGTGGGACACGACGAACAAC 180
Qy 181 AATCTGCCAAAGTATGACTTTAGATCCGACCAATCGAATCGAATCGAATGTAATGG 240
Db 181 AATCTGCCAAAGTATGACTTTAGATCCGACCAATCGAATCGAATCGAATGTAATGG 240
Qy 241 GCGGGGATCTGGAGGGGTTTCGTCAAAAACCTTCTTTATTAACACAGCTGGCGGTAAAG 300
Db 241 GCGGGGATCTGGAGGGGTTTCGTCAAAAACCTTCTTTATTAACACAGCTGGCGGTAAAG 300
Qy 301 ACAATCTGTTTCGCGGTTTGGCAATCTGGATACACTGGCGGACCGGATACAG 360
Db 301 ACAATCTGTTTCGCGGTTTGGCAATCTGGATACACTGGCGGACCGGATACAG 360
Qy 361 GCCTATCAGCGATCTGACGCGGATTTTAAACAGATTGAGAACATTTTCGGGAATTGG 420
Db 361 GCCTATCAGCGATCTGACGCGGATTTTAAACAGATTGAGAACATTTTCGGGAATTGG 420
Qy 421 ACCAATTGACAGTGGTCAATGATGCTCACAAAAGGAATCAAGGTGATTTGGAC 480
Db 421 ACCAATTGACAGTGGTCAATGATGCTCACAAAAGGAATCAAGGTGATTTGGAC 480
Qy 481 TTTCTGCCAATCATTTGACTCTTTTAAAGGCAAGATTCACCTTTGGGAGCGGC 540
Db 481 TTTCTGCCAATCATTTGACTCTTTTAAAGGCAAGATTCACCTTTGGGAGCGGC 540
Qy 541 GCCCTCTACCAATGAACTATATGGCAATTTATTTGATGACGAACAAAGGGTAC 600
Db 541 GCCCTCTACCAATGAACTATATGGCAATTTATTTGATGACGAACAAAGGGTAC 600

Qy 601 TTCCACCATATGGGACATCAGCAACTGGACGACCCGTCAGAGGCCAATGGAATAAC 660
Db 601 TTCCACCATATGGGACATCAGCAACTGGACGACCCGTCAGAGGCCAATGGAATAAC 660
Qy 661 TTCCAGGATCCAGCGGTTTCTCGCTTCCGATTTCTCCAGAAAATGCGACGATGCT 720
Db 661 TTCCAGGATCCAGCGGTTTCTCGCTTCCGATTTCTCCAGAAAATGCGACGATGCT 720
Qy 721 CAATACCTGACCGATCGCGGTTTCAATTTGTAGCACATGAGAGCGATGTTTCGGGAT 780
Db 721 CAATACCTGACCGATCGCGGTTTCAATTTGTAGCACATGAGAGCGATGTTTCGGGAT 780
Qy 781 GATCGGTGAAGCATTTTAAATTCGGGTTTCTCBAATCTGTGCGCGATAAATGTTACCAA 840
Db 781 GATCGGTGAAGCATTTTAAATTCGGGTTTCTCBAATCTGTGCGCGATAAATGTTACCAA 840
Qy 841 AAGAAAGACATTTTCTGTGGGGGATGTTACGGAGATGACCCCGGAAACAGCCAAATCAT 900
Db 841 AAGAAAGACATTTTCTGTGGGGGATGTTACGGAGATGACCCCGGAAACAGCCAAATCAT 900
Qy 901 CTGAAAAGGTCCGGTACGCCAAACACAGCGGTGTCAATGTCTGCTGATTTTGAATCAAC 960
Db 901 CTGAAAAGGTCCGGTACGCCAAACACAGCGGTGTCAATGTCTGCTGATTTTGAATCAAC 960
Qy 961 ACGGTGATTCGAAATGTTTGGGACATTTTACGAAACGATGTACGATCTTAAACAATG 1020
Db 961 ACGGTGATTCGAAATGTTTGGGACATTTTACGAAACGATGTACGATCTTAAACAATG 1020
Qy 1021 GTGAACCAACCGGGAACGAGTACAAATACAAAGAAATCTTAATCACATTTATCGATTAAC 1080
Db 1021 GTGAACCAACCGGGAACGAGTACAAATACAAAGAAATCTTAATCACATTTATCGATTAAC 1080
Qy 1081 CATGATGTCAAGATTTCTTTCGGTAAATTCGAAACAGCGGATTTTGCACAGCGCTT 1140
Db 1081 CATGATGTCAAGATTTCTTTCGGTAAATTCGAAACAGCGGATTTTGCACAGCGCTT 1140
Qy 1141 GCTTTCAATTTCTCTTCGGGGGTACGCCCTCCATCTTATTTGGAACGGAACAATAATG 1200
Db 1141 GCTTTCAATTTCTCTTCGGGGGTACGCCCTCCATCTTATTTGGAACGGAACAATAATG 1200
Qy 1201 GAGCGGGATGACCGGTACAAACCGGGGATGATGCGCGGCTTTGATGACGAAACACC 1260
Db 1201 GAGCGGGATGACCGGTACAAACCGGGGATGATGCGCGGCTTTGATGACGAAACACC 1260
Qy 1261 GCTTTAAAGAGGTGTCAACTCTGCGGGGTTCGCGAGGAACAATGCGCGCATCCAGTAC 1320
Db 1261 GCTTTAAAGAGGTGTCAACTCTGCGGGGTTCGCGAGGAACAATGCGCGCATCCAGTAC 1320
Qy 1321 GGCACCAACCCAGCGTTGGATCAACAATGATTTTACATTTATGAAAGGAAATTTTTC 1380
Db 1321 GGCACCAACCCAGCGTTGGATCAACAATGATTTTACATTTATGAAAGGAAATTTTTC 1380
Qy 1381 AACGATGTGTTGTTGGCCATCAATCGAAACAGCAATCTCTCTATTCGATTCGGGT 1440
Db 1381 AACGATGTGTTGTTGGCCATCAATCGAAACAGCAATCTCTCTATTCGATTCGGGT 1440
Qy 1441 TTGCAAGCGGCTTGGCAATGCGGATTTATCTGTAGGGCTGTTCGGGGGG 1500
Db 1441 TTGCAAGCGGCTTGGCAATGCGGATTTATCTGTAGGGCTGTTCGGGGGG 1500
Qy 1501 AACGGGATTCGGTTTCAATGGAAGTGTGCTTCTGATTCAGCGTTGCGCTTGGAGCGGTG 1560
Db 1501 AACGGGATTCGGTTTCAATGGAAGTGTGCTTCTGATTCAGCGTTGCGCTTGGAGCGGTG 1560
Qy 1561 TCTGTTGGCAGTACAGACATCCGCTTCAGCGCGCAATCGGATCGGTTCGTCGCAAT 1620
Db 1561 TCTGTTGGCAGTACAGACATCCGCTTCAGCGCGCAATCGGATCGGTTCGTCGCAAT 1620
Qy 1621 ATGCGGATTCGGGTAATGTGTCAAGTCAAGCGGAAAGGTTTGGGACGACGAGGA 1680
Db 1621 ATGCGGATTCGGGTAATGTGTCAAGTCAAGCGGAAAGGTTTGGGACGACGAGGA 1680
Qy 1681 ACCGTGACATTTGGCGAGTGCACGAGCTGTGAAATCTTGACACATCCATCGGATTTGAA 1740

```

Db      1681 ACCGTCACATTTGGCGAGTGCACGACCTGTGAATCCCTGGACATCCCAATCGATGGAA 1740
Qy      1741 GTGTACGTTCCCAACATGSCCCCGGCTGACCGATGTGAAAGTCAACCGGGTGGAGTT 1800
Db      1741 GTGTACGTTCCCAACATGSCCCCGGCTGACCGATGTGAAAGTCAACCGGGTGGAGTT 1800
Qy      1801 TCCACCAATCTGTATCTTACCAATATTTGAGTGGACGACGACATCCGTTGCTTTACT 1860
Db      1801 TCCACCAATCTGTATCTTACCAATATTTGAGTGGACGACGACATCCGTTGCTTTACT 1860
Qy      1861 GTGAAAGTGGCCCTCCGACCAACCTGGGGGATGAAGATTACCTGACGGCAACATACCG 1920
Db      1861 GTGAAAGTGGCCCTCCGACCAACCTGGGGGATGAAGATTACCTGACGGCAACATACCG 1920
Qy      1921 GAATTTGGGAATTGGACACGATAGGACGCGCTTAACATCCGACGAGGCCCTG 1980
Db      1921 GAATTTGGGAATTGGACACGATAGGACGCGCTTAACATCCGACGAGGCCCTG 1980
Qy      1981 CTCGGGCCCAATATCCGATTTGGTTTATGTATTCAGCGTTCACGACGGAAGACGATT 2040
Db      1981 CTCGGGCCCAATATCCGATTTGGTTTATGTATTCAGCGTTCACGACGGAAGACGATT 2040
Qy      2041 CAATTCAGATCTTCATCAGCGTCCGATGGAACGATTCATTCGAGGAGATGTTTCAAC 2100
Db      2041 CAATTCAGATCTTCATCAGCGTCCGATGGAACGATTCATTCGAGGAGATGTTTCAAC 2100
Qy      2101 CACGTGGCCCAACTCCCGCGGTGCAACCGGTAACTACTGTTACGTGGCAAACTAG 2160
Db      2101 CACGTGGCCCAACTCCCGCGGTGCAACCGGTAACTACTGTTACGTGGCAAACTAG 2160

```

RESULT 3

```

BACAMYLAA
LOCUS      2726 bp      DNA      linear      BCT 26-APR-1993
DEFINITION B.stearothermophilus maltogenic alpha-amylase (amyM) gene, partial cds.
ACCESSION M36539
VERSION    M36539.1
KEYWORDS   maltogenic alpha-amylase.
SOURCE     Geobacillus stearothermophilus
ORGANISM   Geobacillus stearothermophilus
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
REFERENCE 1 (bases 1 to 2726)
            Diderichsen, B. and Christiansen, L.
            Cloning of maltogenic alpha-amylase from Bacillus
            stearothermophilus
            FEMS Microbiol. Lett. 56, 53-60 (1988)
COMMENT    Original source text: B.stearothermophilus DNA.
FEATURES   Location/Qualifiers
            1..2726
                /organism="Geobacillus stearothermophilus"
                /mol_type="genomic DNA"
                /db_xref="taxon:1422"
                208..2361
                    /note="maltogenic alpha-amylase (amyM)"
                    /codon_start=1
                    /transl_table=11
                    /protein_id="AA022233.1"
                    /db_xref="GI:142501"
                    YQI1IDRFYDGTNNPAKSYGLDPTKSKWYMGDLGVROKLPYLKQLGVTI
                    WLSPLVNNLDTAGTNTGHWYTRDPFKQIEHFGNWTFTLVNDAAHQNKIVTD
                    VYPNHSFPKANDSTABGALYNNGTGMYFDATKGYFHNGDISNWDVRYAQW
                    KNFTDPAGSLADLSQENGTIAQVLTDAVOLVAHLRIDAVKHNSFGSKSLADLY
                    QKQIPLUGVWGDSDGTANLHLEKTRYANNSSUNVLDPLNIVNVEFGTQWDL
                    NNWNQTNKTYKLENLITFDNHMSRFLSVNSKNKALFQRLSPSLRGVRRPIY
                    GTEQYMGAGNDPNRNGMFPATFTTTFKVESTLAGLRNNAIQYGTTRWINDDY
                    YIYERKFPNDVLAIVNRNTQSSYISGLQALPNSYADLSGLGNGISVNSGV
                    ASFTLAPGVSQVQYSTASAPQIGSVAPNMGIPGNVITDGKFGTQGTQVTFPGVT
                    ATVKSNTSNLEVVVPMNAGLTDVKTAVAGVSNLSYVNLISGTQTSVTFVKSAPP
                    TNLGDKIYLQENIPELGNWSTQTSVYNNAGQPLLPAPNPDWFFVFSVPAGKIQKFP
                    FIKRADGTIQENSGNSHNVTPTGTAGNITVTWQV"

```

```

BASE COUNT      736 a      620 c      718 g      652 t
ORIGIN
Query Match      96.7%; Score 2089.6; DB 1; Length 2726;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2146; Conservative 0; Mismatches 4; Indels 14; Gaps 4;
QY      1 ATCAAAAAGAAAACGCTTTCTTTATTTTGGGACCTGATGCTCTCATCGGTCTTCTGTTTC 60
DB      208 ATGAAAAGAAAAGCGCTTTCTTTATTTTGGGACCTGATGCTCTCATCGGTCTTCTGTTTC 267
QY      61 AGCGGTTCTCTTCCTGTAACAATCCAAACCGCGCTGAAGCCAGAGTTCCGCAAGCGTCAAA 120
DB      268 AGCGGTTCTCTTCCTGTAACAATCCAAACCGCGCTGAAGCCAGAGTTCCGCAAGCGTCAAA 327
QY      121 GGGGACGTGATTTACCGAGATTATCATTTGACCGGTTTTACGATGGGACACGACGACAAAC 180
DB      328 GGGGACGTGATTTACCGAGATTATCATTTGACCGGTTTTACGATGGGACACGACGACAAAC 387
QY      181 AATCCTGCCAAAAGTTATGGACTTTACGATCCGACCAAAATCGAAGTGGAATGTATTGG 240
DB      388 AATCCTGCCAAAAGTTATGGACTTTACGATCCGACCAAAATCGAAGTGGAATGTATTGG 447
QY      241 GCGGGGATCTGGAGGGGTTCTGTAAGAACTTCCTTATCTTAAACAGCTGGGCGCTAACG 300
DB      448 GCGGGGATCTGGAGGGGTTCTGTAAGAACTTCCTTATCTTAAACAGCTGGGCGCTAACG 507
QY      301 ACAATCTGGTTCTCCCGGGTTTTGGCAATCTGAGATACACTGGCGGCGACCGATAACACG 360
DB      508 ACAATCTGGTTCTCCCGGGTTTTGAACAATCTGAGATACACTGGCGGCGACCGATAACACG 567
QY      361 GGCATACAGGATCTGACGCGCGGATTTTAAACAGATTGAGGACATTTTCGGGAATTGG 420
DB      568 GGCATACAGGATCTGACGCGCGGATTTTAAACAGATTGAGGACATTTTCGGGAATTGG 627
QY      421 ACCACATTTTACACGCTGGTCAATGATGCTCACAAAACGGAATCAAGGTGATTGTCCAC 480
DB      628 ACCACATTTTACACGCTGGTCAATGATGCTCACAAAACGGAATCAAGGTGATTGTCCAC 687
QY      481 TTGTGCGCAATCATTTGACCTCTTTAAAGCAACGATTCACCTTTGGGGAAGCGGCG 540
DB      688 TTGTGCGCAATCATTTGACCTCTTTAAAGCAACGATTCACCTTTGGGGAAGCGGCG 747
QY      541 GCCTCTTACAAACATGGAACCTATATGGSCAATTTTGTATGACGCAACAAAGGATAC 600
DB      748 GCCTCTTACAAACATGGAACCTATATGGSCAATTTTGTATGACGCAACAAAGGATAC 807
QY      601 TTCCACCAATAATGGGGAACATCAGCAACTGGGACGACCGGTACGAGGCGCAATGGAAAAAC 660
DB      808 TTCCACCAATAATGGGGAACATCAGCAACTGGGACGACCGGTACGAGGCGCAATGGAAAAAC 867
QY      661 TTCCACGATCCAGCGGTTCTCGCTTGGCGGATTTGTGCGAGGAATGCGACGATGCT 720
DB      868 TTCCACGATCCAGCGGTTCTCGCTTGGCGGATTTGTGCGAGGAATGCGACGATGCT 927
QY      721 CAATACCTGACCGGCTTCAATTTGGTAGACATGAGGCGGATGTTTGGCGATT 780
DB      928 CAATACCTGACCGGCTTCAATTTGGTAGACATGAGGCGGATGTTTGGCGATT 978
QY      781 GATCGGTGAGCATTTTAATTCGGGTTCTCCAAATCGTTGGCCGATAAACTGTACCAA 840
DB      979 GATCGGTGAGCATTTTAATTCGGGTTCTCCAAATCGTTGGCCGATAAACTGTACCAA 1038
QY      841 AAGAAAGACATTTTCTCGTGGGGGATGATGACGAGATGACCCCGGAAACAGCAATCAT 900
DB      1039 AAGAAAGACATTTTCTCGTGGGGGATGATGACGAGATGACCCCGGAAACAGCAATCAT 1098
QY      901 CTGAAAGAGTCCGTTACCCCAACACAGCGGTGTCAATGTCTGGATTGTATCTCAAC 960
DB      1099 CTGAAAGAGTCCGTTACCCCAACACAGCGGTGTCAATGTCTGGATTGTATCTCAAC 1158
QY      961 ACGGTGATTGAAATGTCTTCCGACATTTTACGCAACAGATGTACGATCTTACAAATATG 1020

```

Db 1159 ACGGTGATTGCGAAATGTTTCCGACAAATTTACGCAACGATGTACGATCTTAACAATATG 1218
Qy 1021 GTGAACCAAAACGGGGAACGAGTACAAATACAAAGAAATCTAATCAACATTTATCGATAAC 1080
Db 1219 GTGAACCAAAACGGGGAACGAGTACAAATACAAAGAAATCTAATCAACATTTATCGATAAC 1278
Qy 1081 CATGATATGTCAGAAATTTCTTTGGTAAATTC--GAAAGGGGAATTTGACAGGCG 1137
Db 1279 CATGATATGTCAGAAATTTCTTTGGTAAATTTGAAAGAAACAAAGGCGAATTTGACCA-CCG 1337
Qy 1138 CTTGCTTTCAATCTCACTTCGCGGGTACGCCCT--CCATCTATTATGAGAACGCAACATA 1196
Db 1338 CTTGCTTTCAATCTCACTTCGCGGGTACGCCCTCCCACTATTATGAGAACGCAACATA 1397
Qy 1197 CATGGCAGCGCAATGACCGGTACAAACGGGGGATGATCCGCGCTTTGATACGAAAC 1256
Db 1398 CATGGCAGCGCAATGACCGGTACAAACGGGGGATGATCCGCGCTTTGATACGAAAC 1457
Qy 1257 CACGCGCTTTAAGAGGTGTCACTCTGGGGGTGGCAGGAACAATGCGGCATGCA 1316
Db 1458 CACGCGCTTTAAGAGGTGTCACTCTGGGGGTGGCAGGAACAATGCGGCATGCA 1517
Qy 1317 GTACGCAACACACCGCTTGGATCAACAAATGATTTACATTTATGAAACGGAAT 1376
Db 1518 GTACGCAACACACCGCTTGGATCAACAAATGATTTACATTTATGAAACGGAAT 1577
Qy 1377 TTTCAACGATGTCTGTGTGGCCATCAATCGAAACGCAATCTCTCTATTTCGATTTC 1436
Db 1578 TTTCAACGATGTCTGTGTGGCCATCAATCGAAACGCAATCTCTCTATTTCGATTTC 1637
Qy 1437 CGGTTCGACAGCGCTTCCAAATGCGAGTATGCGGATTTCTGTCAGGGGTGTGG 1496
Db 1638 CGGTTCGACAGCGCTTCCAAATGCGAGTATGCGGATTTCTGTCAGGGGTGTGG 1697
Qy 1497 GGGGAACGGGAATTCGGTTTCCAAATGGAAGTGTCTGTTACGCTTGCSCCTGGAGC 1556
Db 1698 GGGGAACGGGAATTCGGTTTCCAAATGGAAGTGTCTGTTACGCTTGCSCCTGGAGC 1757
Qy 1557 CGTGTCTGTGTTGGCAGTACAGCATCCGCTTCAGCGCGCGCAATCGATCGATCGCTGCC 1616
Db 1758 CGTGTCTGTGTTGGCAGTACAGCATCCGCTTCAGCGCGCGCAATCGATCGATCGCTGCC 1817
Qy 1617 AAATATGGGGAATTCGGTATGTTGTTGTCAGATCGACGGAAGAGTTTGGGACGACGCA 1676
Db 1818 AAATATGGGGAATTCGGTATGTTGTTGTCAGATCGACGGAAGAGTTTGGGACGACGCA 1877
Qy 1677 GGGAAACCGTGATTTGGCGAGTACAGCGACTGTGAAATCTCTGACATCCCAATCGAT 1736
Db 1878 GGGAAACCGTGATTTGGCGAGTACAGCGACTGTGAAATCTCTGACATCCCAATCGAT 1937
Qy 1737 TGAAGTGTACGTTCCCAACATGGCGCGGCTGACCGATGTGAAAGTGTACCGCGGCTGG 1796
Db 1938 TGAAGTGTACGTTCCCAACATGGCGCGGCTGACCGATGTGAAAGTGTACCGCGGCTGG 1997
Qy 1797 AGTTTCAGCAATCTGTATTTCTTCAATATTTTGTAGTGGAAACGACGACATCGGTTGTGT 1856
Db 1998 AGTTTCAGCAATCTGTATTTCTTCAATATTTTGTAGTGGAAACGACGACATCGGTTGTGT 2057
Qy 1857 TACTGTGAAAGTGTGCGCTCCGACCACTGGGGATAGATTTACCTGACGGGCAACAT 1916
Db 2058 TACTGTGAAAGTGTGCGCTCCGACCACTGGGGATAGATTTACCTGACGGGCAACAT 2117
Qy 1917 ACCGGAATGGGGAATTTGAGCAGCGATACGAGCGGAGCGGTTAACAAATGCGCAAGGGCC 1976
Db 2118 ACCGGAATGGGGAATTTGAGCAGCGATACGAGCGGAGCGGTTAACAAATGCGCAAGGGCC 2177
Qy 1977 CTTCTCGCGCCCAATTTATCCGGAATGCGTTTATGATTTATTCAGCGTTTCACGAGGAAGAC 2036
Db 2178 CTTCTCGCGCCCAATTTATCCGGAATGCGTTTATGATTTATTCAGCGTTTCACGAGGAAGAC 2237
Qy 2037 GATTCAATTCAGATTTCTTCATCAGCGTGGGATTCGACGATTCATTTGGGAGATGGTTTC 2096
Db 2238 GATTCAATTCAGATTTCTTCATCAGCGTGGGATTCGACGATTCATTTGGGAGATGGTTTC 2297

Qy 2097 GAACCACTGGCCACAACTCCCAACGGGTGCAACCGGTAAACATTAATCTGTACGTGGCAAAA 2156
Db 2298 GAACCACTGGCCACAACTCCCAACGGGTGCAACCGGTAAACATTAATCTGTACGTGGCAAAA 2357
Qy 2157 CTAG 2160
Db 2358 CTAG 2361
RESULT 4
AR139106
LOCUS AR139106 576 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 3 from patent US 6207149.
ACCESSION AR139106
VERSION AR139106.1 GI:14481602
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 576)
AUTHORS Fugisang, C. Crone, and Tsuchiya, R.
TITLE Starch binding domains (SBDs) for oral care products
JOURNAL Patent: US 6207149-A 3 27-MAR-2001;
FEATURES Location/Qualifiers
source
1..576
BASE COUNT 146 a 134 c 166 g 130 t
ORIGIN
Query Match 26.7%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 3.7e-160; Indels 0; Gaps 0;
Matches 576; Conservative 0; Mismatches 0;
Qy 1582 TCCGCTTCAGCGCGCAAAATCGGATCGGTGCTCTCAAATATGGGATTCGGGTAAATGTG 1641
Db 1 TCCGCTTCAGCGCGCAAAATCGGATCGGTGCTCTCAAATATGGGATTCGGGTAAATGTG 60
Qy 1642 GTACGATCGACGGGAAAGGTTTGGGACGACGCGGACCGTGACATTTGGGGAGTG 1701
Db 61 GTACGATCGACGGGAAAGGTTTGGGACGACGCGGACCGTGACATTTGGGGAGTG 120
Qy 1702 ACAGCGACTGTGAAATCTCTGGACATCCCAATCGGATGTAAGTGTACGTTCCCAACATGGCC 1761
Db 121 ACAGCGACTGTGAAATCTCTGGACATCCCAATCGGATGTAAGTGTACGTTCCCAACATGGCC 180
Qy 1762 GCGGGCTGACCGATGTGAAAGTGTACCGCGGGTGGAGTTTCCAGCAATCTGTATCTTAC 1821
Db 181 GCGGGCTGACCGATGTGAAAGTGTACCGCGGGTGGAGTTTCCAGCAATCTGTATCTTAC 240
Qy 1822 AATATTTTGTAGTGAACGACGACATCGGTTGTTTACTGTGAAAGTGTCCGCTCCGACC 1881
Db 241 AATATTTTGTAGTGAACGACGACATCGGTTGTTTACTGTGAAAGTGTCCGCTCCGACC 300
Qy 1882 AACCTGGGGGATAAGATTTTACCTGACGGGCAACATACCGGAATTTGGGAATTTGAGCAAG 1941
Db 301 AACCTGGGGGATAAGATTTTACCTGACGGGCAACATACCGGAATTTGGGAATTTGAGCAAG 360
Qy 1942 GATACGCGGAGCGGTTAAACATGCGAAGGCGCCCTGCTCGCGCCCAATTTATCCGGAT 2001
Db 361 GATACGCGGAGCGGTTAAACATGCGAAGGCGCCCTGCTCGCGCCCAATTTATCCGGAT 420
Qy 2002 TGGTTTTATGTATTTCAGCGCTTCCAGCAGAAAGAGCATTTCAATTTCAATCAAG 2061
Db 421 TGGTTTTATGTATTTCAGCGCTTCCAGCAGAAAGAGCATTTCAATTTCAATCAAG 480
Qy 2062 CTTGCGGATGGAACGATTCATATGGGAGATGTTTCGAAACCACTGGGCCCACTCCCAAG 2121
Db 481 CTTGCGGATGGAACGATTCATATGGGAGATGTTTCGAAACCACTGGGCCCACTCCCAAG 540
Qy 2122 GGTGCAACCGGTACATTAATCTGTACGTGGCAAAAC 2157
Db 541 GGTGCAACCGGTACATTAATCTGTACGTGGCAAAAC 576


```
Db 2135 AATTATGCTGTAAAGTAGCTGGCAATGGCTCAACACCAATGGTATACAACTTACC 2194
QY 1828 TTGAGTGAACCCAGACATCGTGTGTGTTACTGTGAAAAGTGGCGCTCCGACCAACCTG 1887
Db 2195 ATCTTTACGGGTGATCAGGTCACTGTACGATTCGTTATTACAAATGCTTCGACCACTG 2254
QY 1888 GGGGATAAGATTACTCTGAGGGGCAACATACACCGGAATGGGAATTCGAGCAGGATACG 1947
Db 2255 GGTCAAGACATCTATCTGACAGGTACGTGGCAGAACTCGGCAATGGAG-----TAGC 2308
QY 1948 AGCGGAGCGGTTAAACATGCGCAAGGGCCCTGCTCGCGCCCAATATCCGGAATGGTTT 2007
Db 2309 GGCACAAACAGCTATTGGACCTGCATTCATCAGGTATCCACACCTACCACTTGGTAC 2368
QY 2008 TATGATTACAGGTTCCGAGGAGGAAGACAGATCAATCAAGTTCTTCATCAAGCGTGG 2067
Db 2369 TATGATGTGAGGCTGCTCGCGGAAACAACTGGAGTTCAAAATCTTTCAGAAAAAT--- 2425
QY 2068 GATGGAACGATTCAATGGGAGATGTTTGAACCAAGTGGCCACCACTCCACGGGTGCA 2127
Db 2426 GGTGCAACATTTACGTGGGAGAGTGGTTCCAGCCACAGTTCCACCAACCAACCGGGT 2485
QY 2128 ACCGTAACATTTACTGTTACGTGCAAAA 2156
Db 2486 ACAGCCACAGTCAATGTAAACTGCAATA 2514

RESULT 6
LOCUS BCGTASE 2383 bp DNA linear BCT 19-OCT-1992
DEFINITION B.circulans gene for cyclodextrin glycosyltransferase.
ACCESSION X68326
VERSION X68326.1 GI:39419
KEYWORDS cyclodextrin synthase; cyclodextrin-glycosyltransferase.
SOURCE Bacillus circulans
ORGANISM Bacillus circulans
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 2383)
AUTHORS Nitschke, L., Heeger, K., Bender, H. and Schulz, G.E.
TITLE Molecular cloning, nucleotide sequence and expression in
Schierchia coli of the beta-cyclodextrin glycosyltransferase gene
from Bacillus circulans strain no. 8
JOURNAL Appl. Microbiol. Biotechnol. 33 (5), 542-546 (1990)
MEDLINE 91103970
PUBMED 1368573
REFERENCE 2 (bases 1 to 2383)
AUTHORS Schiltz, E.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1992) E. Schiltz, Inst. for Organic Chemistry &
Biochem., Albertstrasse 21, 7800 Freiburg, FRG
FEATURES
source
1..2383
/organism="Bacillus circulans"
/mol_type="genomic DNA"
/db_xref="taxon:1397"
/clone="pBC2"
/clone_lib="lambda EMBL3"
222..2378
/gene="CGTase"
222..2378
/gene="CGTase"
/codon_start=1
/transl_table=1
/product="cyclodextrin glycosyltransferase"
/protein_id="CAA48401.1"
/db_xref="GI:39420"
/db_xref="SWISS-PROT:P50920"
/translation="MFQAKRAFLSTLTLLGLLAGSALPFLPASVAYADPDPTAVTNKQ
SFSTDIYQVFDFLDGNPSNNPTGAAYDATCSNLKLYCGGDMGLINKINDYVFD
LGVTAIVSQVENIFATINYSVNTVAHVWARDPKTNPVFGFMADFNOLITTAH
AKGIVLIDFAPNPTSAMETDTSFAENGRLYDNGTLVGYNDTNGVFNHNGSDFS
SLENGIYNLIDLADFNHNNALIDKFKDAIKLMDMGVDIRVDAVKHMLFGLQKWSN
NSSIYAHKPEVFTGEWFLGSAASADNTDFANKSGMSLLDFRENSAVRNVFRDNTSNM
```

```
mat_peptide
324..2375
/gene="CGTase"
/product="cyclodextrin glycosyltransferase"
BASE COUNT 632 a 633 c 559 g 559 t
ORIGIN
Query Match 17.8%; Score 383.4; DB 1; Length 2383;
Best Local Similarity 53.18; Pred. No. 1.8e-102;
Matches 1112; Conservative 0; Mismatches 906; Indels 75; Gaps 11;
QY 88 GCGCGTGAAGCCAGCAGTTCGCGAAGCGTCAAGGGGACGTGATTTACCAAGATTATCAT 147
Db 330 GACACTGCTGTGACCAACAAACAGAGCTTCAGTACAGATGTGATCTACCAAGTATTACG 389
QY 148 GACCGGTTTACGATGGGACACGACGAAACAAATCCTGCCAAAAGTTATGGACTTAC 207
Db 390 GATCGCTTTTGGACGGCAATCCCTCCAACAAACCCACCG-----GAGCGCTTAT 440
QY 208 GATCCGACCAATCGAAGTGGAAATGTATTGGGCGGGGATCTGGAGGGGTTCTGCAA 267
Db 441 GATGCCACATGCGCAATTTGAAGCTGTACTGCGGCGGAGACTGCGCAAGGTTAATCAAC 500
QY 268 AAACCT-----TCCTATCTTAAACAGCTGGCGGTAAACGACAACTCTGTTGTCCCGGTT 321
Db 501 AAAATCAACGATAACTATTTCAGTGATCTGGGTGTCAACGCTTGTGGATCTCCAGCCT 560
QY 322 TTGGCAATCTGGATACACTG-----GCGGCGACCGATTAACACGCGGTATCACGGA 372
Db 561 GTCGAAATATTTTTCGACGATCAACTATAGCGGTGTAACCAATACTGCCTTACCACGGC 620
QY 373 TACTCGAGCGCGATTTTAAACAGATTGAGGAAACATTTCCGGGATTTGGACCACTTGAC 432
Db 621 TACTGGGCTCGAGACTTCAAAAGAACCAATCCGTACTTCGGCAGCATGGCGACTTTCAG 680
QY 433 AGTGTGTCATGATGTCACCAAAACGGAATCAAGTGATTTGTCACCTTTTGCCCCAAT 492
Db 681 AACCTGATTACGCGCTCATGCCAAGGCATCAAGATTGTCATTGACTTTGCACCGAT 740
QY 493 CATTCGACTCCTTTAAGCAACCAATTCACCTTTCGGGAAGCGCGCCCTCTTACAC 552
Db 741 CACACCTCTCCAGCGATGGAAACCGACACCTCTTTTCGGAATAATGGCAGACTGTACGAT 800
QY 553 AATGGAACCTATATGGGCAATTTTGTGATGACGCAACAAAGGGTACTTCCACCAAT 612
Db 801 AACGTCACCTGGTAGCGGGTTATACCAATGA---TACCAGCGGGTACTTCCATCAAT 857
QY 613 GGGGATCAGCAACTGGGACGACCGGTACGAGCGCAATGGAAAAAATTCACCGATCCA 672
Db 858 GCGGCTCCGACTCT-----CTTCCCTGGAGAACGGCATCTACAA 899
QY 673 GCGGTTTCTCGCTTCGCGATTGTCGAGGAAATGGCAGGATTCCTCAATACCTGACC 732
Db 900 AACCTGTATGACCTGGCGGCACTTCAACACCAATTAATCGACCATCGACAAATACTTCAA 959
QY 733 GATCGCGGGTTCAATTTGGTAGCACATGAGCGGATGTTTGGCGGATTTGATCGGTGAAG 792
Db 960 GATCGGATCAACTGTGGCTTGATATGGCGTTGACGGTATTCGGGTGATCGGTGAA 1019
QY 793 CATTTTAATTCGGGTTCTCCAAATCGTTGGCCGATTAATCTGTACCAAAAGAAACAT 852
Db 1020 CATATGCTCTCGGTTGGGAAAAGAGCTGATGTCTCCATCTACGCACACAAACCCGTG 1079
QY 853 TTCTGTGGGGGTAATGGTACGGAGATGCCCGGGAACACCAATCATCTGGAAGAGTC 912
Db 1080 TTCACCTTTTGAGAAATGGTTCTTTGGGATCAGCTCATCCG-----ATGCAGATAACACG 1133
```

Qy 913 CGGTACGACCAACACAGCGGTGTCATATGCTGCTGATTTGATCTCAACACGGTGATTCGA 972
 Db 1134 GATTTTGTCTACAAAGTCTGTAAGCGCTGCTCACTTCGCTTTAACTCTCGGTGCGT 1193
 Qy 973 AATGTTTCGGCAACATTTACGCAAGCATGTACGATCTTAAACAATATGTCGAACCAACG 1032
 Db 1194 AATGTTTCGGTACACACAGCTCAACATGTACGCTCTGATTTCCATGATTAACAGACA 1253
 Qy 1033 GGGACGAGTACAAATACAAAGAAATCTAATCACTTTATGATATGATACCAATGATATGCA 1092
 Db 1254 GTTACGGATCAACACCAAGTGAACGATCAGGTGACGTTCAACCAACATGATGATGAT 1313
 Qy 1093 AGATTTCTTTTCGGTAAATTCGAACAAAGCGAATTTGACCAAGCGCTTGCCTTTCATTTCTC 1152
 Db 1314 CGTTTCAAAAACAGTGGCGTCAACAATCGCGCTCTGGAACAGGCTTTGCGCTTCACATTCG 1373
 Qy 1153 ACTTCGGGGTACGCCCTCACTCTATTATGGAACCGAACAATACATGATGCGAGCGGCAAT 1212
 Db 1374 ACTTCACGTGTGTACCTGCGCATCTACTATGCTACTGAACAGATTTTACGCGGGAATGGC 1433
 Qy 1213 GACCGGTAAACCGGGGATGATCGCGGCTTGTGATACGAACACACCGCTTTAAAGAG 1272
 Db 1434 GATCAGATACCGGGCCAAATGCGTCTGCTTCCAAATCCACTACACTTTTACGTC 1493
 Qy 1273 GTGTCAACTCTGGCGGGTTCGCGAGGAACAATGCGGGATCCAGTACGCGACACACACC 1332
 Db 1494 ATAAGTAAACTCGCGGCTCTCGCAATCCAAATCGCGCATTCGCTACGCTTCCACACAG 1553
 Qy 1333 CAGCGTGGATCAACAATGATGTTTACATTTATGAAGGAAATTTTCAACGATGCTGTCG 1392
 Db 1554 CAGCGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1613
 Qy 1393 TTGGTGGCCATCAATCGAAACACGCAATCTCTTATTCGATTTCCGCTTTGCGACGCGC 1452
 Db 1614 GTTGTGCGGTAAACCGCAATCTTCCAGCTCTGCAAGCATTAACAGGTTAAGCACTTCC 1673
 Qy 1453 TTGCAATGCGAGTATGCGATTTATCTGTACGGGCTGTTGGGGGGACCGGATTTCC 1512
 Db 1674 CTGCTTACAGGCTCATACAGGATGCTGTGGCGGGTGTGAACGGAATTAACATCAC 1733
 Qy 1513 GTTTTCAATGGAAGTGTGCTTGGTTCAGCTTGGCGCTGCGCTGGAGCGGTGTCTTTTGGCAG 1572
 Db 1734 TCCACGAATGTTAGATTAACAATTCACCTTGTCTGGGGTGCACGCGCATGTGGCA 1793
 Qy 1573 TACAGCACTTGGCTTACGCGCGCAATCGGATCGTGTCTCAATATGGGATTCG 1632
 Db 1794 TACACAACTGGCGAAACACACCAACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1853
 Qy 1633 GGTAAATGCTCAGATCGACGGGAAAGGTTTGGGACGACGCGAGGAACCGTGACATTT 1692
 Db 1854 GGTAAATGCTCAGATTTGATGGCGGTGATTCGCTGCGCTGCGAGGACGCGCTACTTC 1913
 Qy 1693 GGC-----GGAGTGAACAGGACTGTGTAATCTCTGCAATCCATCGATCGATTTGAAGTG 1743
 Db 1914 GGCACATACAGCGGTTTACCGGAGCGCGATTAAGTCTTGGGAAGTACACAGATTAAGTA 1973
 Qy 1744 TAGTTTCCCAACATGCGCGCGCGGCTGACCGATGTGAAAGTACCGCGGGTGGAGTTCC 1803
 Db 1974 ACCATCCCTTCGTTGCGCGCTGCGCACTATGCAATGCAAGTTCGCGGACGCGGGTAAAC 2033
 Qy 1804 AGCAATCTCT---AFTCTTCAATATTTTGAAGTGAACGCAATCGGTTGTGTTTACT 1860
 Db 2034 AGCAATGCAACAACAATTCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2093
 Qy 1861 GTGAAGTGGCGCTTCCACCACTTGGGGATAGATTTACTTACCGCGGCACTACCG 1920
 Db 2094 GTAAACAATGCG---TCCACAACGCTTGGACAGAACCTTATTTGACAGCAACGTAGCC 2150
 Qy 1921 GAATTTGGGAATGAGACGCAATGATGACGCGGAGCGGTTTAAATGCGCAAGGCGCCCTG 1980
 Db 2151 GAGCTTGGCACTGAGACACCGGTTCA-----CTGCCATGCAACCGCATTCATCAG 2204
 Qy 1981 CTCGCGCCCAATATTCGGAATGCTGTTTATGATGATTCAGGTTCCAGGAGGAAGACGAT 2040

Db 2205 GTCAATCATCAATACCAACCTGGTACTATGATGTGTCAGGTACCGGACGCAACAGCTG 2264
 Qy 2041 CATTCAAGTCTTCAATCAAGCGTGGGATGGAACGATTCATGGGAGATGGTTTGAAC 2100
 Db 2265 GAGTTCAAAATTTT---TCAAGAAAAACGGTTCACGATTCATATGGGAAACGGTTCTAAC 2321
 Qy 2101 CAGTGGCCACAACTCCACAGGGTGCACGCGGTACATTAATCTGTTACGTGGCA 2153
 Db 2322 CACACATTCATCAACAGAGCGGGAACAGCCCGTTACGTTACGTTGCACTGGCA 2374

RESULT 7
 LOCUS BLCTAG 2516 bp DNA linear BCT 12-SEP-1993
 DEFINITION Bacillus licheniformis cgtA gene for cyclomaltodextrin
 glucanotransferase (CGTase) (EC 2.4.1.19).
 ACCESSION X15752
 VERSION X15752.1 GI:39565
 KEYWORDS cgtA gene; CGTase; cyclomaltodextrin glucanotransferase;
 glucanotransferase; transferase.
 SOURCE Bacillus licheniformis
 ORGANISM Bacillus licheniformis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE 1
 AUTHORS Hill, D.E., Aldape, R. and Rozzell, J.D.
 TITLE Nucleotide sequence of a cyclodextrin glucosyltransferase gene,
 cgtA, from Bacillus licheniformis
 JOURNAL Nucleic Acids Res. 18 (1), 199 (1990)
 MEDLINE 90174931
 PUBMED 2137908
 REFERENCE 2 (bases 1 to 2516)
 AUTHORS Hill, D.E.
 TITLE Direct Submission
 JOURNAL Submitted (04-JUL-1989) Hill D.E., Genetics Institute, Inc., 87
 Cambridgepark Drive, Cambridge, MA 021140, USA

FEATURES
 source
 1..2516
 /organism="Bacillus licheniformis"
 /mol_type="genomic DNA"
 /db_xref="taxon:1402"
 /clone_lib="pUC19 derivative"
 210..2366
 /note="cyclomaltodextrin glucanotransferase (AA 1-718)"
 /codon_start=1
 /trans_cable=11
 /protein_id="CAA33763.1"
 /db_xref="GI:39566"
 /translation="MFQAKRVLLSTLTFTSLLAGSALPFLPASAIYADADTAVTNKO
 NSTDVLVQVPTDRFLDGNBNNTGAAFDGTCNLKLYCGDMOGLVKNKINDVYFSD
 LGVLAWISQVENVLPSTINYSQVNTAYHGYWARDPKKNPYFGMTDFQNLVYTAH
 AKGKILIDFAPNTPSPMEIDTSPFAENGKLDNGNLVGGTINDTNGYFHHNGGSDFS
 TLNGYIKLXLDLADLNNSTIDTFFKDAIKLMDMGVDGIRDAVHMPQGWQKRW
 MSYIAHPFTFTEGFLGAPADNDTDFANESGSLIDFENSARVNRVNDNTSM
 YALDSMLPATADYNQVNDQVTFIDHMDRFTSAVNNRLEQALAFITSRGVPAI
 YVTEOYLTGNGDQDNGKMPSPKTAENVISKLAPLKNPAIAYGSTQORWINN
 DVIVYERKGSVAVAENVNLTPTSTINLTNSIPSGTYDVLGGVLENNITSSGG
 NISSFTLAGATAVQWQTASSETPTTGHVPGVMPKPNVITIDRGFGSAKTVIFGT
 TAVTGATTSWEDTQIKVTIPPVAGGVAVKVAANGVNSNAYINDFTILSGDQSVRFV
 INNAATYAGENIYLTGNVSELGNWTTGAISTGAPFNQVTHAYPTWYTYDVSVPAGKLE
 FEFKNGANITLWEGSNHTFTTSGTATVTINQ"

BASE COUNT 647 a 673 c 602 g 594 t
 ORIGIN

Query Match 17.3%; Score 374.2; DB 1; Length 2516;
 Best Local Similarity 52.3%; Pred. No. 1e-99;
 Matches 1078; Conservative 0; Mismatches 928; Indels 57; Gaps 9;

Qy 112 AGCTCAAGGGGACGTGATTACAGATTATCATTCAGCCGTTTACGTGGGACAG 171
 Db 342 AACTTCAGTACGAGTGTCAUTATCAAGTATTTTACGACCCGATTTCTGGATGTAACCA 401

172 ACAGACAACTCTCGCCAAAGTTATGAGCTTACGATCCGACCAATCGAGTGGAA 231
Db
402 TC---CAACAACTCTACCGGGCTGCCTTTGATGCGAGCTGCACAACTGAGCTGTAC 458
Qy
232 ATGATATGGGGCGGGGATCTGAGGGGGTTCGTCAAAACTTCCTTATCTTAAACAGCTG 291
Db
459 TGGGTGGGAGCTGCGAGGGATTGGTCAACAAAATCAATGACAATATTTCAGTGACTTG 518
Qy
292 GGGGTAACACAACTCTGTTGCTCCCGGTTTGGHCAA---TCTGGATACATCGGGGGC 348
Db
519 GGTGTACAGGCCCTCTGATCTCCAGCCTGTGCAAAATATTTCGCTCACTCAACTAC 578
Qy
349 ACCGAT-----AACACGGGTATACGGATATCGACGCGGATTTTAAACAGATTGAG 402
Db
579 AGGGGTGTAACCAATACAGCTATACGGCTATTGGCACCGGATTTCAAGAGACCAAT 638
Qy
403 GAACATTTGGGAAATGGACCAATTTGACAGTTGGTCAATGATGTCAACAAAACGGA 462
Db
639 CCGTATTTGGGAACCATGACCGATTTTCAGAACTCTGGTAACACCGGCCATGCAAGGC 698
Qy
463 ATCAAGGTGATTTGCACTTTGTCGCCATCATTCGACTCCTTTTAAAGCAACGATTCC 522
Db
699 ATCAAAATCATTTATGATTTGCGGCCAAACCATACGTCCCTGCCATGGAACCGATACC 758
Qy
523 ACCTTTGCGGAAGGGCGGCCCTCTCAACAATGGAACTTATATGGGCAATATTATTGAT 582
Db
759 TCCTTCGCCGAGAACGCAACTGTAGATTAACGCAACTGTGTTGGGGGTACCAAT 818
Qy
583 GAGCGAAACAAAGGTACTTCCACCATAATGGGGACATCAACAATCTGGACACCGGTAC 642
Db
819 GA---TACCACCGGATATTCCACCAATGGCGGCTCGGATTCT----- 861
Qy
643 GAGGCGCAATGAAAAAATCTCAACGATCCAGCGGTTCTCGCTTCGCGGATTTGTCGAG 702
Db
862 ----CCACTCTTGAGATGGGATTTTCAAAAACCTCTAGATCTGGCGGCTCTGATAC 917
Qy
703 GAAATATGGACGATTTGCTCAATACCTGACCGATGGGGGTTCAATTTGGTAGCACATGA 762
Db
918 AATAACAGCAAGATGATACATATTCAAGACGCCATCAAGCTGTGGCTTGATGGGC 977
Qy
763 GCGGATGTTTCCGATTTGATGCGGTGAGAGCATTTTAATTCGGGGTTCTCCAAATCGTG 822
Db
978 GTGACGCGATTTCTGTGTCGATCGGCTCAAGCACATGCTCAGGCTGGCAGAGAAGCTGG 1037
Qy
823 GCGGATAAATCTTACCAAAAGAAAGACATTTCTCTGTGGGGAACTGTACGGAGATGAC 882
Db
1038 ATGTCTATCTATCTGACACAGCGGTATTTCCTTCGGGGAATGGTTCTTGGGATCT 1097
Qy
883 CCGGAAACAGCCATCATCTGGAAGAGTCCGGTACGCCAACACACAGCGGTGTCAATGTG 942
Db
1098 GCTGCACCGG-----ATGCGGATAACACAGATTTTGGCAATGATCTGGTATGATTG 1151
Qy
943 CTGGATTTGATCTCAACAGCGTGAATTCGAAATGTTTCGGCAATTTAGCGCAACGATG 1002
Db
1152 CTGATTTCCGTTTCAATTCGGCTGTCGCAACGTTTCGGGATAACACTTCCAACTG 1211
Qy
1003 TAGCATCTTAAACAATATGTTGAACAAAGGGGAGAGTACAAATACAAAGAAATCTA 1062
Db
1212 TAGCGCTGATTTCCATCTGAGGCTACGGCAGAGATTAATCATCAGTGTGATGACAA 1271
Qy
1063 ATCAATTTATCATTAACATGATATGTCAAGATTTCTTCGGTAAATTCGACAAAGGG 1122
Db
1272 GTCACTTTTATGACAAACATGATGGACCGTTTCAAAACGAGTGTGTAACACCGC 1331
Qy
1123 AATTTCACAGCGGCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTAT 1182
Db
1332 GGTCTGGAACAGGCTCTGCGCTTCACTGCTTCACTGCTTCACTGCTTCACTTATAT 1391
Qy
1183 GGAACCGAACAATATATGCGAGGGGCAATGACCCGTACAAACCGGGGAGTATGCGGGC 1242
Db
1392 GGTACCGAGCATATCTGATGGAACGGGACCGGATACCGGGGCAAAATGCTCTCC 1451
Qy
1243 TTTGATACGAAACACCGCTTTTAAAGAGGTGTCAATCTTGGCGGGGTTGGCAGGAAC 1302

1452 TTCTCCAAATCGACTACAGGTTCAAGTGATCAGAGCTGGCACCTCTGCGAAATCC 1511
Qy
1303 AATGGCGGATCCAGTACGGCACCCACCGAGTTTGGATCAACATGATGATTTTACATT 1362
Db
1512 AACCGCGGATTCGCTACGGTTCCACACAGCAGCGCTGGATCAACATGATGATATATC 1571
Qy
1363 TATGAACGGAATTTTCAACGATGTCTGTGTGGTGGCCATCAATCGAAACACGCAATCC 1422
Db
1572 TATGAGCCAAAGTTTGGCAAAAGCGTTCCGCTGCTTCCGTTAACCGCAATCTCACAGC 1631
Qy
1423 TCTTATTCGATTTCCGTTTTCAGACGCGCTTGCACAAATGCGAGCTATCGGATTTCTG 1482
Db
1632 CCAACCATGATCACTAACCTGATACGCTTCCATCAGAACCTTACCGATGCTC 1691
Qy
1483 TACGGCTGTTTGGGGGGAAACGGGATTTCCGTTTCCAATGGAAGTGTCTGTTCTGTTACG 1542
Db
1692 GCGCGGTGCTGAACGGAACCAACATCACGTCAGTGGAGGCAATATTCTTCTTCAAG 1751
Qy
1543 CTTGCGCTCGAGCGGTGTCTGTTTGGCAGTACAGCACTCCGCTTCAGCGCGGCAATC 1602
Db
1752 CTCGCTGAGGAGCTACCGCTGTTTGGCAATATACGGCAAGTGAACACGACGCCAACCATC 1811
Qy
1603 GGATCGGTGCTCCAAATATGGGGATTCGGGTAAATGTGTACGATCGACGGGAAGGT 1662
Db
1812 GGTACAGTTGGCCCTGTAAATGGGTAAAACAGGTAAACGTCTTACCATCAGCGACGGG 1871
Qy
1663 TTTGGGACGACGACGGAACCGGTGACATTTGGCGAGTGAACGACTG-----TG 1713
Db
1872 TTCGTTCCGCCAAAGGTACCGCTACTTCTGCTCAACAGCGCTTACGGGCTCAGCCATT 1931
Qy
1714 AATCTCTGCAATCCATCGGATTTGAGTGAATGTCCTCCAAATGCGCGCGCGGTGACC 1773
Db
1932 ACTTATGGGAGACACTCAGATCAAGTCAACATCCACAGTAGCAGCGGTGACTAT 1991
Qy
1774 GATGTAAAGTCAACGCGGTGGAGTTTCCAGCAATCTGTATTCTTCAATATTTTGAAT 1833
Db
1992 SCAGTGAAGTAGTSCCAATGTTGTHACAGCAACCGCTTAAACGATTTACCATCTT 2051
Qy
1834 GGAACGACAGATCGGTTGTGTTTACTGTGAAAAGTGGGCTCCGACCAACCTGGGGAT 1893
Db
2052 AGCGCGATCAGGTATCGGTGGGTTCTGTCATCAATATGCCACAACCTGCGTGGCGAG 2111
Qy
1894 AAGATTTACCTGACGCGCAACATACCGGAATTTGGGATTTGGAGCAGATACGAGCGGA 1953
Db
2112 AATATCTACCTGACAGGAAACGTTCCGAACTCGGTAACTGGACCAAGTGC-----A 2165
Qy
1954 GCGGTAAACATGCGCAAGGGCCCTGCTCGGCCCAATTTATCCGATTTGTTTATGTA 2013
Db
2166 GCTTCCATTTGACCGGCTTCAATCAGGTTCATCCAGCTTACCGGACTTGGTATTATGAC 2225
Qy
2014 TTCAGCGTTCCAGCAGGAAGAGCATTCATTTCAAGTCTTCTCATCAGCGTGGATGGA 2073
Db
2226 GTAAGTGTTCGGCGGGGAAAACAGCTGGAAITCAAGTTCTTCAAGAAAACCGGCT--- 2282
Qy
2074 ACGATTTCAATGGGAGATGTTTCGAACCAACCGCGGCCCAACTCCACCGGTCAACCGGT 2133
Db
2283 ACCATTACGTGGGAGGTGATCCCAATCACACCTTTACACACCGAGAGTGTACTGCC 2342
Qy
2134 AACATTACTGTTCGTGGCAAAA 2156
Db
2343 ACAGTAACGATAAACTGGCAATA 2365

RESULT 8

BSCGIDNA

LOCUS

DEFINITION

Bacillus sp. cgt gene for beta-cyclodextrin glucanotransferase.

ACCESSION

X66106

VERSION

X66106.1 GI:39838

KEYWORDS

cyclodextrin glucanotransferase; cyclomaltoextrin

Bacillus sp.

BSCGIDNA

Bacillus sp. cgt gene for beta-cyclodextrin glucanotransferase.

ACCESSION

X66106

VERSION

X66106.1 GI:39838

KEYWORDS

cyclodextrin glucanotransferase; cyclomaltoextrin

Bacillus sp.

ORGANISM Bacillus sp.
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE 1. (bases 1 to 2489)
 AUTHORS Akhmetzjanov, A.A.
 JOURNAL Unpublished
 REFERENCE 2. (bases 1 to 2489)
 AUTHORS Akhmetzjanov, A.A.
 TITLE Direct Submission
 JOURNAL Submitted (21-MAY-1992) A.A. Akhmetzjanov, Russian Academy of Sciences, Bashkirian Scientific Centre, Dept of Biochemistry & Cytochemistry, Pr Oktjabrja 69, Ufa 450054, USSR
 FEATURES Location/Qualifiers
 1. 2489
 /organism="Bacillus sp."
 /mol_type="genomic DNA"
 /strain="6.6.3 (ACM-1.7.9.3.-D)"
 /db_xref="taxon:1409"
 /clone_lib="pBCGT 8"
 /clone_id="pBCGT 8"
 212..2368
 /gene="cgt"
 212..2368
 /gene="cgt"
 /EC_number="2.4.1.19"
 /codon_start=1
 /transl_table=11
 /product="cyclomaltodextrin glucanotransferase"
 /protein_id="CAA46901.1"
 /db_xref="GI:39839"
 /db_xref="SWISS-PROT:P31747"
 /translations="MFQMKARFLSTLTILGLLAGSALPFLPASAAVADPDIAVTHKQ
 SFSFDVIYQFTDFRFLDNPSPNTGAAYDATCSNLKYCGGDWGLINKINDNYPD
 LGVPLAISQPVNIFATINSYNTAYHYWARDFKTNPFYGTWADFQNLITFAH
 AKGKI1IDPAPNHPSPAMETDTSFAENGKLYDNLGVGTNDYNTDGYFHHNGSDFS
 SLENGIYKLDLPAPNHNATIDKYFKDAIKLMDGVDRVDAVKHIALGWQKSW
 MSSLYVHKPFTGFEFLGSAADANTDFPANKSGLLDREFNSAVNFVFRONTNSM
 YALMSINSTADYDQNDQVTFIDNMDRPKTSVAVNRRLLEQALAFILISGVPL
 YVGEQTLTNDYDNRKWPFSKSTTFNFWISKLAPKNSFALAYGTQORWLN
 DVIYVEFKSVAVVAVNRNLSPANTGLSTSLPTGSYDVLGVNGNNTISNG
 SVNGFILAAGATVAVWQYAAETPTTIGHVGVNPKGNVYVTDGRGFGTGTVYFV
 TAVTGAATISWEDQIKVTIPSAAGNVAVKVAANGVNSVNHFTILTDQVTVFV
 INNAATVLGNIYVLTGNVAELGNSTGTSTALGPAFQVTHQVPTWYVDVSPAGKLE
 PKFKKNGSTIHWEGSNHKTTPASFAIVTNWQ"
 BASE COUNT 683 a 637 c 571 g 598 t
 ORIGIN
 Query Match 16.8%; Score 363.4; DB 1; Length 2489;
 Best Local Similarity 51.7%; Pred. No. 1.7e-96;
 Matches 1096; Conservative 0; Mismatches 956; Indels 69; Gaps 9;
 QY 64 GGTCTCTTCGTACATCAACGCCGCTGAAGCCAGAGTTCGCCAGAGTCAAGGG 123
 DB 296 GCTTCTGTGATACGCCGACCGGACATGCTGTCAACCAACAAAGCTTCAGTACA 355
 QY 124 GACGTGATTACCAAGATTATCATTCACCGGTTTACCATGGGACACGACCAACAAT 183
 DB 356 GATGTAATCTACCAAGTGTTCACAGACCGCTTTTTCGCGCAATCCCTCCACAA 411
 QY 184 CTGGCAGAAAGTTATGGACTTTACGATTCGACCAAAATCGAAGTGGAAATGTTGGGC 243
 DB 412 -----TCCAAACAGGAGCGGCTTATGATGCGACATGTAGCAATCTGAAACTGTATGCGSGT 466
 QY 244 GGGGATCTGGAGGGGTTCTGCAAAAAC-----TCCTATCTTAAACAGCTGGGGCTA 297
 DB 467 GGAGACTGGCAGAGTTAATCAAAAATCAATGATTAATTTCACTGATCTGGGTGC 526
 QY 298 ACAGCAATCTGTTGTCCCGGTTTGGCAATCTGGATAACCTGGCGGCAACGA----- 353
 DB 527 ACAGCGTGTGTGATCTCCAGCGCTGTGAAATAATTTTTCGACGATCAACTACAGCGGT 586
 QY 354 -----TAACACGGGCTATCAGGATCTGACGCGCGATTTTAAACAGATTGAGGAACAT 408
 DB 587 GTAAACCAATACGGCTACCATGGCTACTGGCCCGGGACTTCAAAAAGACAAATCCGTAC 646

409 TTCGGAAATTGGACCAACATTTGACAGTTGTGTCATGATGTCACCAAAACGGAATCAAG 468
 DB 647 TTCGGACAAATGGCCGAGCTTCCAAACCTGATGATGAGGGCCCATGCCAAGGGCATCAAG 706
 QY 469 GTGATTGTGCACTTTGTGTCCTCAATCATTCGATCCTTTTAAAGCAAAACGATTCACCTTT 528
 DB 707 ATTATTATTGACTTTGGCCCGAACACACCTCTCTGCGCATGGAACGACACTTCTTTT 766
 QY 529 GCGAAGGGCGGCGCTCTACACATGGAACCTATATGGGCAATTTATTTCATGACGCA 588
 DB 767 GCGAGAAATGGAATAATTTGATAGTACGCAACATGTTAGGCGGTTCACATAATGAT--- 823
 QY 589 ACAAAGGGTACTTCCACCATAATGGGACATCAACAACTGGGACGACCGTACGAGCG 648
 DB 824 ACCAATGGCTATTTCCATCAACAGCGGCTTCCGACTTCTCTTC----- 868
 QY 649 CAATGAAAAATTCACGGATCCAGCGGTTTCTGCTGTCGCCGATTTGTTCGAGGAAAT 708
 DB 869 ---CTGGAGACGGCATCTACAAAAACTTTGTATGACCTGGCGGATTTCAACCAACAC 925
 QY 709 GGCACGATTGCTCAATACCTGACCGGATCGCGGTTCAATTGGTGTAGCACATGAGCGAT 768
 DB 926 GCAACCATCGACAGTACTTCAAGACGCCATTAAACTGTGGCTCGATATGGGCGTTGAC 985
 QY 769 GGTTCGGATGTATGCGGTGAAGCAATTTTAATTCGGGGTTTCTCAAAATCGTTGGCCGAT 828
 DB 986 GGCATTCGGGTAGATGCGGTGAAGCATATCGCTCTCGGCTGGCAAAAGAGCTGGATGCT 1045
 QY 829 AAACCTCCAAAGAAAGACATTTCTCTGGTGGGGAATGTTACGGAGATGACCCCGGA 888
 DB 1046 TCCATTTAGTACAAACCGGTTTCACTTTTGTGATGGTTCCTGG-----GATCA 1099
 QY 889 ACAGCCAATCATCTGGAAGAAGTCCGCTAGCCCAACACAGCGGTCTCAATGTGTGGAT 948
 DB 1100 GCTGCTCTGATCGAGATAACACGGATTTTCTACAAAGCTGCTGATGAGCCTGCTCGAC 1159
 QY 949 TTTGATCTCAACCGGTGATTCGAATGTGTTCGGCAATTTACGCAACGATGTACGAT 1008
 DB 1160 TTCGGTTTAACTCTGCGGTGCGTAATGTTGTTCCGTGATAACACTTCAACATGTACGCT 1219
 QY 1009 CTTAAACAAATATGCTGAACAAACGCGGAACGAGTACAATAACAAAGAAATCTTAATCA 1068
 DB 1220 CTGGATTCCATGATTAACAGTACCGCTACGATTAACACCAAGTCAATGATCAGTGA 1279
 QY 1069 TTTATCGATACATGATATGTCAAGATTTCTTCGCTAAATTCGAACAGCGCAATTTG 1128
 DB 1280 TTCATCGACATCATGATGATGATGCTTTCAAAAACAGTGGGTCAACAATCGCGCGCTG 1339
 QY 1129 CACGAGCGCTTGTCTTCAATTTCTCACTTCGCGGGTACGCCCTCCATCTATTATGGAACC 1188
 DB 1340 GAGCAGGCTCTGCGCTTTCAGTGAATTCACGTGTGTGCTGCCATTTACTACGGCA 1399
 QY 1189 GAACAATAATACGAGCGGCAATGACCCGTACACCGGGGGATGATGCGGGGTTTGTAT 1248
 DB 1400 GAACAGTATCTACAGGTAAATGTGATACCGGATAACCGAGCAAAATGCTTCTGTTCTCC 1459
 QY 1249 ACACAAACCAACCGCTTTAAAGAGGTGTCACTCTGGCGGGTTGCGCAGGACAAATGCG 1308
 DB 1460 AAATCCCAACACCGCTTTTACGTCATTAAGCAAACTGGCACTCTGGGCAAAATCCCAATCCG 1519
 QY 1309 GCGATCCAGTAGGCAACCAACCCAGCGTTGGATCAACAATGATGTTTACATTTATGAA 1368
 DB 1520 GCGATTGCTACGCTCCACGACGAGCGCTGGATTAACATGATGTTTATGTTATGAG 1579
 QY 1369 CGGAATTTTTCAGATGTGTTGTTGGTGGCCATCAATCGAAGACGCAATCTCTCTAT 1428
 DB 1580 CGAAAAATTTGGCAGAGGTAGCGGTGTTGTTGGGTAAACCGCAATCTGTCCACACGCG 1639
 QY 1429 TCGATTTCGGTTTTCAGACGGCTTGCCAATGCACATGCGACTATGCGGATTTATCTGTACGG 1488
 DB 1640 AATATTACAGGGCTGAGCACTTCTCTGCTACAGCTCATACACGGATGTCTCTTGGCGGT 1699

QY 1489 CTCTTGGGGGGAACGGGATTCGGTTTCCAATGGAAGTGCCTTGGTTCAAGCTTCGG 1548
 Db |||
 QY 1700 GTACTGAACGGAACAACATCAGTCCAGTAATGGCAGGTTACTTTCCTTCCCTTGTCT 1759
 Db |||
 QY 1549 CTGGAGCCGCTCTGTGTTGGCAGTACAGCATCCGTTTCAAGCGCGCAATTCGGATCG 1608
 Db |||
 QY 1760 GCGGGTGCAACGGCAGTATGCGATATACAGCAGCGCAACACCAACCAACCGGACAT 1819
 Db |||
 QY 1609 GTTGCTCCAAATCGGGATTCGGGTATGTTGCTCAGATCGACGCGGAAGGTTTGGG 1668
 Db |||
 QY 1820 GTGGTCCGGTCTATGGCAACACGAGCAATGTTGTCAGCATGATGTCGCGGATTCGGC 1879
 Db |||
 QY 1669 ACACCGACGGAACCGTCACATTTGGC-----GGAGTGAACGCACTGTGAATCC 1719
 Db |||
 QY 1880 TCCACTAAAGCAGCGTATCTTCGGCACTACAGCTGTTCGGGAGCAGGATCACTCT 1939
 Db |||
 QY 1720 TGACATCCAAATCGGATGGAAGTGTAGTTTCCACATGCGCGCGCGGCTGACCGATGTC 1779
 Db |||
 QY 1940 TGGGAAGATACACAGATTAAGATTAACGATTCATCCATCCGTGGTGTGCTGTAATGCTGTC 1999
 Db |||
 QY 1780 AAAGTCACCGCGGGTGGAGTTTCCAGCAATCTGTATTCTTACAATATTTTGTGTCGAACG 1839
 Db |||
 QY 2000 AAAGTAGCTGCGAAGCGGGTAAACAGCAATGCATATAACCAATTCACATCTCGACCGGC 2059
 Db |||
 QY 1840 CAGACATCGGTTGTGTTTACTGTGAAAGTGCGCCTCCGACCAACCTGGGGGATAGATT 1899
 Db |||
 QY 2060 GATCAGTTACCGTAGCGCTTGTCTATTACAAATGCTTCCACACGCTTGACAGACATA 2119
 Db |||
 QY 1900 TACTGACGGGCAACATACCGGAATGCGGAATTTGGACACGATACGACGGGAGCCGTT 1959
 Db |||
 QY 2120 TATCTGACAGGCAACAGTAGCTGAGCTTGGCAATGCGCACTGGACAGAGTTCCA-----CTGCC 2173
 Db |||
 QY 1960 AACAAATGCGCAAGGCGCCCTCTCGGCCCAATATCCGATTTGGTTTATGATTCAGC 2019
 Db |||
 QY 2174 ATTGGACCAAGCAATTCATCAGTGATTCATCATACCACTGGTACTATGATGTCAGC 2233
 Db |||
 QY 2020 GTTCCAGCAGGAAGACGATTCAATTCAGATTTCTTCATCAAGCGCTGGGATGGAACGATT 2079
 Db |||
 QY 2234 GTACCGGCGAGTAAAGACTTGATTCAGTTCT--TCAAGAAAACGGTTCACGATT 2290
 Db |||
 QY 2080 CAATGGGAATGTTTGGACACGAGTGGCCACACATCCCAAGGTGCAACCGGTACATTT 2139
 Db |||
 QY 2291 ACATGGGAAGCGGTTCTTAACACACAGTTCCACACACCAAGCAAGCGGAACAGCAAGGTT 2350
 Db |||
 QY 2140 ACTGTTTACGTGCAAAACTG 2160
 Db |||
 QY 2351 ACGTGAACTGGCAGTAAGAG 2371
 Db |||
 RESULT 9
 BC251CGT BC251CGT 2197 bp DNA linear BCT 14-SEP-1994
 LOCUS B. circulans strain 251 cgt gene.
 DEFINITION
 ACCESSION X78145
 X78145.1 Gi:510491
 cgt gene; cyclodextrin-glycosyltransferase.
 KEYWORDS
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 ORGANISM
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE
 1 Lawson, C.L., van Montfort, R., Strokopytov, B., Rozeboom, H.J.,
 Kalk, K.H., de Vries, G.B., Penninga, D., Dijkhuizen, L. and
 Dijkstra, B.W.
 Nucleotide sequence and X-ray structure of cyclodextrin
 glycosyltransferase from *Bacillus circulans* strain 251 in a
 maltose-dependent crystal form
 J. Mol. Biol. 236 (2), 590-600 (1994)
 MEDLINE 94149761
 PUBMED 8107143
 REFERENCE 2 (bases 1 to 2197)
 AUTHORS Penninga, D.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-1994) D. Penninga, Microbial Physiology,

University of Groningen, Biologisch Centrum, P O Box 9750 AA Haren,
 NETHERLANDS
 FEATURES
 source
 1..2197
 /location/Qualifiers
 /organism="Bacillus circulans"
 /mol_type="genomic DNA"
 /strain="251"
 /db_xref="taxon:1397"
 45..2197
 /gene="cgt"
 45..49
 /gene="cgt"
 56..2197
 /gene="cgt"
 /EC_number="2.4.1.19"
 /codon_start=1
 /transl_table=11
 /product="cyclomalto-dextrin glucanotransferase"
 /protein_id="CAA55023.1"
 /db_xref="GI:510492"
 /db_xref="SWISS-PROT:P43379"
 /translation="MKKFLKSTAAALALGLSLTGLFSPAQAAPDTSVNKNQFSTDI
 YQIFTRPSDGNPANNPTCAAFDGTCLNRLCYGQWQGLINKINDVLTGWTALW
 ISQVENIYGIINYSVANTAYHGVARDPKTNPAYGTIADFNLIARAAHKIKVI
 IDPAENITSPASSDQPSFAENGRLVDNGLTGGVTDNTQNLPHNGGTFDFSTAEVLY
 KNLVDLADLANHNSSTVDVVLKDALKMWLDLGDIGIDMDAVKMPGCKQKSPAAVNY
 KPVTFGEWFLGVNEVSPENHKPANESGMSLLDFRFAQKRVQFRDNTDNYLKAHL
 EGSAAIDYQVDDQVTFIDNHMERPHASNRRLKLBALAFITLSRGVPAIYGTQY
 MGGTDPDRNRPISFTSTAYQVTKLAPLRKCNPAIAYSTQERINNDVLIYER
 KFSGNVAVAVANENANAPASISGLVTSIQGSYNDVILGGLLNGTLISVSGGAASNT
 LAAGTAWOYTAATATPTIGHVGPMAKPGVTITIDGRFGSSKGTVVFGTAVSQA
 DITSNEDTOIKVPIPAVAGGNVNIKVAAGTASNYVDNFEVLSDGVSVFVFNNAI
 TALQNYILGSGVSELGNWDPAKALGPNINQVVIQPNWYIDVSPAGKTIEFLFKK
 QGSTVTEGSGSNHTFTAPSSGTATINWQD"
 56..130
 /gene="cgt"
 131..2194
 /gene="cgt"
 /product="cyclomalto-dextrin glucanotransferase"
 128..133
 /gene="cgt"
 /note="signal processing site gcgcgc"
 BASE COUNT 528 a 659 c 606 g 404 t
 ORIGIN
 Query Match 16.3%; Score 353; DB 1; Length 2197;
 Best Local Similarity 52.1%; Pred. No. 2.1e-93;
 Matches 1080; Conservative 0; Mismatches 915; Indels 78; Gaps 10;
 QY 111 AAGCGTCAAAGGACGCTGATTACAGATTTATCATTCAGCGGTTTACGATGGGACAC 170
 Db |||
 QY 166 AATTTTCAGCAGCGATCTATCAAAATTTTCCAGCAGGTTTTCGACGGCAATCC 225
 Db |||
 QY 171 GAGCAACAAATCTCCCAAAAGTTATGACATTCGATCCGACCAAAATCGAAGTGA 230
 Db |||
 QY 226 CGCAACAATCCGACCGC-----GCGCGGTTTGAACGACCTGCACGACCTCG 276
 Db |||
 QY 231 AATGATTGGCGCGGATCTGAGGGGGTTCGCAAAACT-----TCCTATCTTAA 284
 Db |||
 QY 277 GCTGATTTCGCGCGCGCATCTGCGAGGGCATCATCAACAAATCAACGCGGTTCCTGAC 336
 Db |||
 QY 285 ACAGCTGGCGCTAAACGCAATCTGGTGTTCGCCGCTTTTGACAACTGATACACT--- 341
 Db |||
 QY 337 CGGATGGCGGTTACCGCATCTGGATCTCCAGCGGTGAAACATCTACAGATCAT 396
 Db |||
 QY 342 -----GCGCGGCAACGATCAACCGGCTATCAGGATCTCGACGCGGATTTAAACA 395
 Db |||
 QY 397 CAATTATTCCGGCGTAAACCAACGCGCTTATCAGGCTTCTGGCCCGCGGACTTCAAGAA 456
 Db |||
 QY 396 GATTGAGGAACATTTCCGGAATGGACCACTTTGACACGTTGGTCAATGATGCTCACA 455
 Db |||
 QY 457 GACGAATCCGCCCTACGGCAGCATTCGGACTTCAGAACTTCGACCGCGCGCATGC 516

QY 456 AAACGAATCAAGTGAATGTCGACTTTGTGCCCAATCATTCGACTCTCTTTTAAGGCATA 515
 Db 517 AAAAAACATCAAGTCAATTCGACTTTGTGCCCAATCATTCGACTCTCTTTTAAGGCATA 576
 QY 516 CGATTCCACCTTTGGGAGGCGCGCCCTCTACACATGGAACCTATATGGGCAATTA 575
 Db 577 CCAGCTTCTCTTTCGGAACCGCGCTGTACGATAAAGCGACGCTGCTCGGGGATA 636
 QY 576 TTTTGTATGACGCAACAAAGGCTACTCCACCAATATGGGACATCAGCAACTGGGACGA 635
 Db 637 CACGAAAGATACGAGAACTGTTCCACATAA-----CGGCG 675
 QY 636 CGGTAAGAGCGCAATGGAATAAATTACAGGATCGAGCGGTTCTCGCTTGGCGATT 695
 Db 676 CACGACTTTTCCACACCGAAGAAAGGATCTACAAAACTGTACGATCTCGCGGACCT 735
 QY 696 GTCCGAGGAATGCGACAGTTCCTCAATACCTGACCGATGGCGGTTCAATTTGGTAGC 755
 Db 736 GAACCTAACACAGACCGTGGAGCTCTTCTGAGGACCGCATCAAAATGTGCTGGA 795
 QY 756 ACATGAGCGGATGTTTGGGATTTGATCGGTGAAGCAATTTAAATCGGGTTCTCCAA 815
 Db 796 CTTGGGATCGAGCGATCGCATGATGATGCGGTGAAGCATATGCCGTTCCGCTGGCAGAA 855
 QY 816 ATCGTTGGCCGATTAAGTGTACCAAAAGAAAGCAATTTCTGCTGGTGGGGAATGTACGG 875
 Db 856 GAGCTTTATGCTGCGGTCAACACTATAAGCCGGTCTTTACCTTCGCGCAATGTTCTCT 915
 QY 876 AGATGACCCCGGACAGCAATCATCTGAAAAGGTCGGGTACGCCCAACACAGCGGTGT 935
 Db 916 GGG-----CGTAAATGAAGTGAGCGCGGAAACCAATAGTTTTCGCAAGATCGGCAT 969
 QY 936 CAAATGCTGATTTGATCTCAACACGCTGATTCGAAATGCTTCCGCAATTTAGCA 995
 Db 970 GAGCTGCTTCAATTCGCTTTTCCCAAAAGGTGCGCAGGTGTTCCGCGGCAACACCGA 1029
 QY 996 AACGATGATACCATCTTAACATATGTGTGAACCAAGCGGACGAGTACAAATCAAGA 1055
 Db 1030 CAATATGATACCGCTTGAGCGGATGCTGGAGGCTCCGACGCTATGCGCCAGGTGGA 1089
 QY 1056 AAATCTAATCAATTTATGATTAACATGATATGTCAAGATTTCTTTTGGTAAATTCGAA 1115
 Db 1090 TGACCAAGTGAGCTTCAATCAACCAACCATGATGAGCGGTTTCCACGCAAGCAATGCAAA 1149
 QY 1116 CAGCGAATTTGACAGCGCTGCTTTCATCTCTCACTTCGCGGGTACGCCCTCAT 1175
 Db 1150 CCGCGGAGCTGAGCAGCGCTTCCGTTCACTGCTCGCGGCTCCCGCAT 1209
 QY 1176 CTATTATGAAACCAATACATGCGAGCGGCAATGACCCGTATGAGTCAACCAATGAT 1235
 Db 1210 TTATTACGCAACCGGACGATGATGTCGGGCGGACCGATCCGGAACAACCGGCGCGAT 1269
 QY 1236 CCGCGGTTTATGACGACACCAACCGCTTTAAAGAGTGTCACTCTGCGGGGTTGCG 1295
 Db 1270 CCGCTTCTTCTCAGCTGAGCAGCGCTTATGAGTCAATTTCAAAAGCTGCGCGCTGCG 1329
 QY 1296 CAGGAACATGCGCGATTCAGTACGCAACCAACCGCGCTGAGTCAACCAATGAT 1355
 Db 1330 CAAAGTCAACCGGCGATTCAGTACGCAACCGGCGCTGATCAACCAACGAGT 1389
 QY 1356 TTACATTTATGAAAGGAAATTTTCAACGATGTGTGTTGGTGGCCATCAATCGAAACAC 1415
 Db 1390 GCTCAATTTATGAGCGCAATTCGCGACGCAACGTTGCGGTGTTGCGCTCAACCGCAAT 1449
 QY 1416 GCATCTCTCTATTCGATTTCCGTTTTCAGACGCGCTTGCCTAAATGCGAGTATGCGGA 1475
 Db 1450 AAACGCGCGCTTCCATTTTCGCGGACTTGTCTCTTCTTCCCTGCGCAAGGAGCTACAGGA 1509
 QY 1476 TTATCTGTCAGGCTGTTGGGGGGGAAACGGAATTTTCGTTTCCAAATGGAAGTGTGCTTC 1535
 Db 1510 CGTCTTGGCGGCTTCTGAACGCAACACGTTATCGGTAGGCTCCGCGGGCGGCGCTC 1569
 QY 1536 G----TTACGCTTGGCGCTGAGCGGCTGTCTGTTTGGCGATACAGCATCGCTTACG 1592

Db 1570 CAATTTCACTGCTTGGCGCGGCGGACGCGGTGTGGCAGTACACCGCGCTTACGGCGAC 1629
 QY 1593 GCGCAATTCGATCGGTCTCTCCAAATATATGGGATTCGCGGTAATGTGTCACGATCGA 1652
 Db 1630 GCGCACTTGGCGCATGTCGGGCGGATGATGCGCAAGCGGCGGTGACGATCAGATCGA 1689
 QY 1653 CCGGAAAGTCTTGGGACGACGAGGAAACGTCGATGATTTTGGCGGAGTGACAGCGACTG- 1711
 Db 1690 CCGCGCGGCTTGGCTCTAGCAAGGCAACCGTCTACTTCGGTACGACGCGGTGAGCGG 1749
 QY 1712 -----TGAATCTGCGACATCCGATCGGATTTGAAGTGTAGCTTCCCAACATCGCGC 1763
 Db 1750 GCGCGACATTCAGCTCTTGGGAAGACGACGAGATCAAGTGAATAATTCGCGCGCTGCGAGG 1809
 QY 1764 CCGGCTGACCGATGTGAAGATC---ACCGCGGTTGGAGTTTCCAGCAATCTGATTTCTTA 1820
 Db 1810 CCGCACTTACATTAAGTTCGCAACGCTCGCGAAGCGCAAGCAATGTGATGACAA 1869
 QY 1821 CAATTTTGTAGTGAAGCAGACATCGGTTGTGTTTACTGTGAAAAGTGGCCTCCGAC 1880
 Db 1870 CTTGAGGTATTGTCGCGAGACGAGGTGAGGTCGCTTCTGTCGTCACCAACGCGACGAC 1929
 QY 1881 CAACCTGCGGATTAAGATTTTACCTGACGCGCAACATACCGCAATTTGGGAATTTGAGACAC 1940
 Db 1930 GCGCTTGGGCAAAATGTGATCTGACGCGGAGTGTGAGGAGCTGGGAGTGGGACCC 1989
 QY 1941 GATACGAGCGGAGCGGTAAATGCGCAAGGCGCCCTGCTCGCGCCCAATATCCGGA 2000
 Db 1990 G-----GCAAAAGCAATCGGCGGATGTACATCAGTGTGTTTACCAATATCCGAA 2040
 QY 2001 TTGCTTTTATGATTCAGCGTTCCAGCAGAAAGACGATTCATTTCAAGTCTTTCATCA 2060
 Db 2041 CTGATTTATGATGCTGAGGTTTCGCGCGCAAGATGAGTTCAGTTCAGTTTGGAAA 2100
 QY 2061 CGCTGCGGATGGAACGATTCATTTGGGAGAAATGTTTCGAACGAGTGGCCCAACTCCAC 2120
 Db 2101 AC---AAGCTCCACGTCAGTGGGAAAGCGGCGAGCAACCACTTCACCGCGCGCTC 2157
 QY 2121 GGTGCAACCGTAACTACTTGTAGTGGCA 2153
 Db 2158 CAGCGCACCGGACCATTAACGTGAATGGCA 2190

RESULT 10
 E01516
 LOCUS E01516 2058 bp DNA linear PAT 29-SEP-1997
 DEFINITION DNA encoding cyclomal-dextrin glucanotransferase.
 ACCESSION E01516
 VERSION E01516.1 GI:2169772
 KEYWORDS JP 1988039597-A/3.
 SOURCE Paenibacillus macerans
 ORGANISM Paenibacillus macerans
 Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
 REFERENCE 1 (bases 1 to 2058)
 AUTHORS Sugimoto, T., Kubota, M. and Sakai, S.
 TITLE SUGAR TRANSFORM REACTION AND PRODUCTION OF FOOD AND DRINK OBTAINED BY USING SAID REACTION
 JOURNAL Patent: JP 1988039597-A 3 20-FEB-1988;
 HAYASHIBARA BIOCHEM LAB INC
 COMMENT OS Bacillus macerans
 PN JP 1988039597-A/3
 PD 20-FEB-1988
 PF 30-MAR-1988 JP 1986072290
 PI SUGIMOTO TOSHIYUKI, KUBOTA MICHIO, SAKAI SHUZO PC
 C12P19/18, C12N9/10, C12N15/00/A23L1/22, A23L1/236, (C12P19/18, PC
 C12R1:07);
 PC (C12N9/10, C12R1:07), (C12N15/00, C12R1:07);
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: clone-pMAU210;

FH Key Location/Qualifiers
 FH mat_peptide 1..2058 /product='cyclomaltdextrin
 FT glucanotransferase'
 FT Location/Qualifiers
 1..2058 /organism='Paenibacillus macerans'
 /mol_type='genomic DNA'
 /db_xref='taxon:4252'

FEATURES

source

BASE COUNT 517 a 634 c 524 g 383 t
 ORIGIN

Query Match 16.1%; Score 348.8; DB 6; Length 2058;
 Best Local Similarity 52.3%; Pred. No. 3.8e-92;
 Matches 1078; Conservative 0; Mismatches 912; Indels 72; Gaps 11;

QY 123 GGAGCTGATTAACAGATATCATTTGACCGGTTTACGATGGGGACAGCAACAA 182
 DB 42 GGTACGGTTTACAGATTGTAACCGACCGGTTTGTGGACG--CAATTCCGCCACAA 98
 QY 183 TCCTGCCAAAGTATAGACTTACGATCCGACCAATCGAATCGAATGGAATGTAATGGG 242
 DB 99 CCGACCGGAGAGCGCTTCAGCAGCGATCATTCCAACCTGAAGCTGTATTTCGGGGCGA 158
 QY 243 CGGGATCTGGAGGGGTTCTGTCAAAACCTTCCTTATCTTAAACAGCTGGGCGTAAGC 302
 DB 159 CTGCGAGGATCAGCAAAATCAGCGGCTATCTGACGAATGGGATCACCGC 218
 QY 303 AATCTGGTTTCCCGGTTTGGACAATCTGGATCACT-----GGCGGCAACGA 353
 DB 219 CCTCTGATCTCGACGGTTTGAGAACATCACCGCGTCATCAATTATTCGGCGTCAA 278
 QY 354 TAAACGGGTATCAGGATCTGACCGGGATTTAAACAGATGAGGAACAATTCGG 413
 DB 279 CAATACAGCTTACACGGTTACTGGCTTCGCGACTTCAAGAAAGCAATGCGCGTTCCG 338
 QY 414 GAATTGACCAATTTGACAGCTTGGTCAATGATCTCACCAAGCGAATCAAGTGT 473
 DB 339 CAGCTTCACGACTTCTCCAAATTGATGCGGAGCGCATTCACCAATATCAAGGTAGT 398
 QY 474 TGTGCTTTGTGCCCAATCAATGACTTCCTTTTAAAGCAACAGATTCACCTTTGGGA 533
 DB 399 TATGACTTTGACCTTAATCACCAACCGGGTTGAGTACGACCCCTCGTTTGGCCGA 458
 QY 534 AGGGCGCCCTCTACCAATGCAATGCACTATATGGCAATTAATTTGATGACGCAACA 593
 DB 459 GAACGGCGCTCTACCAACGGAAAGCTGCTCGGCAAGTA---TAGCAACGATACCGG 515
 QY 594 AGGGTACTTCCACATAATGGGGACATCAGCACTGGGACGACCGGTACGAGGCGCAATG 653
 DB 516 CGGCTGTTCCACCAATGGCGGACCGATTCT-----CGAGAC 557
 QY 654 GAAAAATTTCAGGATCCAGCGGTTCTCGCTGCGATTTGTCCGAGAAATGGCAC 713
 DB 558 TGAAGCGGTATCTACAGAACCTGTACGATCTCGGGATATCAATCAGAACACAAAC 617
 QY 714 GATTGCTCATATCTGACGATGCGGCGGTTCATTTGATGACATGAGAGCGGATGGTT 773
 DB 618 CATGACTGCTATCTCAAGAAATGATCCAGCTGTGGCTGAATCTCGGAGTCCAGCGCAT 677
 QY 774 CGCGATTGATCGGCTGAAGCATTTAATTCGGGGTTCTCAAAATCGTTGGCCGATAAAT 833
 DB 678 CGGCTTCGACGCGTGAAGCATATGCTCAGGCGTGGCAGAGACTTACGTTCTCGTCA 737
 QY 834 GTACCAAA----AGAAAGCAATTTCTGTTGGGGGAATGTGACGAGATGACCCCGAAC 890
 DB 738 CTACAGCAGCGCAATTCGCGTTTCACTTCGTTGTAATGTTTCTCGGCCCGCGAGAAAT 797
 QY 891 AGCCAATCATCTGAAAAGGTCGGTACGCGCCCAACACAGGGTGTCAATGCTGATTT 950
 DB 798 GACC-----CAGCAACATCACTTCGGAATCAGCGGCGATGCACCTGCTGGACTT 851

QY 951 TGATCTCAACACGGTGATTTCGAATGTGTTCCGACATTTAGCGAAAGATGTACGATCT 1010
 DB 852 TGGTTTGGCGAGAAATCCGTTGAAGTGTTCGCGCAAGATCGGAGACGATGACCGACT 911
 QY 1011 TAACAATATGTTGTAACCAACCGGGAACGAGTACAAATACAAAGAAATCTAATCAATT 1070
 DB 912 GAACTCGGTGATCTCCAGCACCGGCTCCAGCTATATATACATCAACACATGTTAGCT 971
 QY 1071 TATCGATAACCATGATATGTCAGATTTCTTTCGGTAAATTCGAACAGGGAATTTGA 1130
 DB 972 CATCGAAACCATGATGATGACCGCTTCAGCAAGCCGAGCGAGCACTGCCGCCGCA 1031
 QY 1131 CCAGGCGTTGCTTTCACTTCTCGGSGTACGCCCTCCATCTATATATGAAACGA 1190
 DB 1032 GCAGGCTTTCGGTAAACGCTGACTTCCCGCGCGTTCGGCAATCTACTACGGTACAG 1091
 QY 1191 ACAATCATGGCAGCGCAATGACCCGTACAAACCGGGGATGATCGCGGTTTGATAC 1250
 DB 1092 GCAATATATGACCGCAACGGGACCGCAACCGCGCATGATGACCGGTTTCGATAC 1151
 QY 1251 GACAAACACCGCTTTAAAGAGGTGCACTCTGGGGGGTTGCGAGGAACAATGCGGC 1310
 DB 1152 GAAACAAGACGGTACAAAGTATCAAGGCGCTGGCTTCGGCTTCGCAAGTCCAAACCGG 1211
 QY 1311 GATCCAGTACGCAACACACCGAGGTTGGATCAACAAATGATGTTTACATTTATGAAG 1370
 DB 1212 TCTGCGCTACGCTCGACGCCAGCGTTGGTGAACAGCGACGTTACGTATATGACG 1271
 QY 1371 GAAATTTTCAACGATGTCGTTGTTGGCCATCAATCGAAACACGCAATCTCTCTATTTC 1430
 DB 1272 CAAGTTTCGGAAGCAACGTAGCTCTCGTTGCGTCAACCGAGCTCGACGATGCTATTC 1331
 QY 1431 GATTTCGGTTTTCAGAGCGCTTCCAAATGCGAGCTATGCGGATTTATCTGACGGGT 1490
 DB 1332 GATATCGGAGCGCTTACTGCTCTGCCAAAGGAACTATACCGACGTTCTCGCGGCT 1391
 QY 1491 GTTGGGCGGAAACGGATTTCCGTTTCCAAATGAGTGTGCTTCGTTTCGCTTCGGCC 1550
 DB 1392 GCTTATGCGAATTCAAATTCGTTTACGGCGCACGTTTACAGCACTTTACACTTGAGC 1451
 QY 1551 TGGAGCGGTGTCTGTTTGGCAGTACAGCACATCCGCTTCAGCGCGCAAAATCGGATCG 1610
 DB 1452 GGGCGGTACGGCAGTCTGGCAGTACAGCAGCGGAACTCTCGCGATTATTCGCAAGT 1511
 QY 1611 TGCTCCAAATATGGGATTCGGGTATGTTGTCAGTACGACGGAAGAGGTTTGGGAC 1670
 DB 1512 CGGCGGACTATGGCAAGCGCGGCAACACCATCAGTACGACGCGCGGCTTCGGTAC 1571
 QY 1671 GACGCAAGCAACCGTGACATTTGG-----CGGAGTGACAGCGACTGTGAAATCCTG 1721
 DB 1572 TAGAAGAACAAAGTTACTTTTCGTAGCAGCGCTTACCGGCGGAACATCTGTAGCTG 1631
 QY 1722 GACATCCAAATCGGATTTGAAGTGTACGTTCCCAATGCGCGCGCGGCTGACCGATGTGA 1781
 DB 1632 GGAAGTACCGAAATCAAGGTCAAAGTTCGAAAGTGGCGCGCGCAACACGCGCGTTAC 1691
 QY 1782 AGTCACGCGGCTGGAGTTTCCAGCAATCTGTATTTCTTACAAATTT-----TTGAGTG 1835
 DB 1592 GGTAAAGCAACCGCGCGGCACTACAGCGCGGTTTACAACTTTAACGTTACTGACTGC 1751
 QY 1836 AACGCAACATCGGTTGTGTTTACTGTGAAAGTGGCTTCGACCAACCTCGGGGGATAA 1895
 DB 1752 CGATCAGGTCACTGTCCGCTTCAAAGTCAACATGCCACACGCG---CCTGGGACAAA 1808
 QY 1896 GATTTCCTGACGGCAACATACCGGAATTTGGGAAATTTGGAGACGGATACGAGCGGAGC 1955
 DB 1809 CGTCTACCTGACCGTAACTGCGCGAGGTTTGGCAACTGG-----ACAGCGCCCAA 1859
 QY 1956 CGTTTAAATGCGAAGGGCCCTGCTCGGCCCAATTTACGATTTGTTTATTTATTT 2015
 DB 1860 CGCAATCGGTCCGATGTACAAACAGGTAGAGCGAGTATCCGACTTTGGTACTTCGACGT 1919
 QY 2016 CAGCGTTCCAGCAGGAAAGACGATTCAATTTCAAGTTCTTCATCAAGCGTTCGGATGGAAC 2075


```

Db 1920 CAGCGTTCCGGCCCAACACACGGCGCTGCAATTCAAGTTTCATCAAGTGAACGGCTGACAGT 1979
Qy 2076 GATTCATATGGAGAGATGGTTCCAAACAGTGGCCACACATCCCAAGGTTGCAACCGGTAA 2135
Db 1980 GACT---TGGGAGGGCGGCAACACACACCTTCACTCGCTTCGAGCGCGGTGGAC 2036
Qy 2136 CATTAAGTGTACGTGGCRAAAC 2157
Db 2037 CGTACAGTGTGAGCAGAC 2058

RESULT 11
E00874
LOCUS      2139 bp      DNA      linear      PAT 29-SEP-1997
DEFINITION DNA encoding cyclomultodextrin glucanotransferase.
ACCESSION E00874
VERSION    1
KEYWORDS  JP 1986135581-A/2.
SOURCE    Paenibacillus macerans
ORGANISM  Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
REFERENCE 1 (bases 1 to 2139)
AUTHORS   Sugimoto T., Kubota M. and Sakai S.
TITLE     POLYPEPTIDE HAVING CYCLOMALTODEXTRIN GLUCANOTRANSFERASE ACTIVITY
JOURNAL   HAYASHIBARA BIOCHEM LAB INC
COMMENT   OS Bacillus macerans
          PN JP 1986135581-A/2
          PD 23-JUN-1986
          PF 14-OCT-1985 JP 1985228169
          PI SUGIMOTO TOSHIOYUKI KUBOTA MICHIO, SAKAI SHUZO PC
          CI 2129/10, CI2129/00, (CI2129/10, CI2129/07);
          CC strandedness: Double;
          CC topology: Linear;
          CC hypothetical: No;
          CC anti-sense: No;
          CC *source: strains=17A;
          CC *source: clone=pXAU 210;
          FH Key Location/Qualifiers
          FT sig_peptide 1..81
          FT CDS 1..2141
          FT /product='cyclomultodextrin
          FT glucanotransferase'.
          FT Location/Qualifiers
          FT 1..2139
          FT /organism='Paenibacillus macerans'
          FT /mol_type='genomic DNA'
          FT /db_xref='taxon:44252'

BASE COUNT 535 a 652 c 552 g 400 t
ORIGIN

Query March 16.1%; Score 348.8; DB 6; Length 2139;
Best Local Similarity 52.3%; Pred. No. 3.8e-92;
Matches 1078; Conservative 0; Mismatches 912; Indels 72; Gaps 11;

Qy 123 GGACGTGATTTACAGATTTATCATTTGACCGGTTTACGATGGGACACAGACAAACAA 182
Db 123 GGATACGTTTACAGATTTGATACCGCGGTTTGTGACGG---CAATTCCGCAACAA 179
Qy 183 TCTCGCAAAAGTTATGCACTTTACGATCCGACCAAAATCGAATGAAATGATTTGGG 242
Db 180 CCCGACCGGACGCGCTTCAGCAGCGATCATTTCCAACTGATTTTGGGGGCGA 239
Qy 243 CGGGGATCTGAGGGGTTGTCAAAACCTTCTTACTTTAAACAGCTGGGCTAACAC 302
Db 240 CTGGCAGGGGATTCACGAACAAATCAACGACGCGCTATCTGACCGAATGGGATCACCGC 299
Qy 303 AATCTGGTTGTCGGGTTTGGACAAATCTGGATACACT-----GGCGGGCACCGA 353
Db 300 CTTCTGGATCTCGACGCGGTTGAGAACATCACCGCGCTCATCAATTATTTCGGGGCTCA 359

```

```

Qy 354 TAACACGGCGTATCACGGATATGACGCGGATTTTAAACAGATTGAGGAACATTTGG 413
Db 360 CAATACAGCTTACACAGGTTTACTGGCTTCGGACTTCAAGAGACCAATGCGCGGTTGG 419
Qy 414 GAATGGACACATTTGACAGCTTGTGATGATGCTCACAAACAGGAATCAAGGTGAT 473
Db 420 CAGCTTCAACCGACTTCTCCAAATTTGATGCGCGACGCGATTCACAAATATCAAGGTAGT 479
Qy 474 TGTGCACTTTGTGCCCAATCATTTCGACTCTTTTAAAGCAACAGATTCCACCTTTGGCGA 533
Db 480 TATGACATTTGCACTAATCAGACCAACCGCGTTCAGTACGACCCCTCTGTCGCGA 539
Qy 534 AGCGGCGGCTTCAACCAATGGAACCTTATATGCGCAATTTATTTGATGACGCAACAAA 593
Db 540 GAACGGCGGCTTCAACCAACGGAACGCTGTCGGCAAGTA---TAGCAACGATACCGC 596
Qy 594 AGGTTACTTCCACCATATGCGACATCAGCACTGCGGACGACCGGTACGAGGCGCAATG 653
Db 597 CGGCTGTGTTCCACCAATGCGGACCGGATTTCT-----CGACGAC 638
Qy 654 GAAAACTTACGGATCCAGCGGTTTCTCGCTTCCGATTTTGTGCGAGGAAAAATGGCAC 713
Db 639 TGAAGCGGTATACAGAACCTGTACGATCTCGGATATCAATCAGAACACAAACAC 698
Qy 714 GATGCTCATACCTGACCGATCGCGGTTCAATGTTGTTAGCACATGAGGCGGATGGTTT 773
Db 699 CATCGACTGTATCTCAAGGAATCGATCAGCTGTGGCTGAATCTCGGAGTTCGACGGCAT 758
Qy 774 GCGGATTGATGCGGTGAAGCATTTTAAATTCGGGGTTCTCCAAATCGTTGGCGGATAAACT 833
Db 759 CGGCTTCGACCGCTGAAGCATATGCTTCAGGCTGGCGAGAGAGCTACTCTCTCGTCAAT 818
Qy 834 GTACAAA---AGAAAGACATTTTCTGTTGGGGAATGTTAGGAGATGACCCCGGAAC 890
Db 819 CTACAGACGCGCAATCCGGTGTTCACCTTCGGTGAATGGTTCTCGGCGCCGAGAAAT 878
Qy 891 AGCAATCATCTGGAAGAGTCCGGTACCCCAACAGAGGTTCAATGTGCTGATTT 950
Db 879 GAC-----CAGGACACATCACTTCGGAATCAGAGGCGCATGCACCTGCTGATTT 932
Qy 951 TGATCTCAACACGGTGAATCGAAATGTGTTGGCACATTTACGCAACAGATGTACGATCT 1010
Db 933 TGGCTTTGCGCAGGAAATCCGTTGAAGTGTTCGCGACAAAGTCGGAGACGATGACCGA 992
Qy 1011 TAACATATGTTGAACCAACAGGGAACAGGTACAAATACAAAGAAATCTAATCAAT 1070
Db 993 GAATCGGTGATCTCAGCACCGGCTCCAGCTATTAATTAATCAATCAATGATGTTACGTT 1052
Qy 1071 TATCGATAACCATGATATGTCAAGATTTCTTTTCGGTAAATTCGAACAGGCGAATTTGCA 1130
Db 1053 CATCGAACACCATGATGACACCGCTTCCAGCAAGCGGAGCGAGCACTCGCGCGACCGA 1112
Qy 1131 CGAGGCGCTTCTTTCATTTCTCACTTTCGGGGTACGCCCTCCATCTATTTATGGAACCGA 1190
Db 1113 GCAGGCTCTTGGGTGATGATGCTTCCCGCGGCTTCGGCAATCTACTACGATACAGA 1172
Qy 1191 ACAATACATGGCAGCGGCAATGACCCGTTACCAACCGGGGATGATCGCGCTTTGATAC 1250
Db 1173 GCATATATGACCGGCAACGGCGACCCGACACACCGCGCATGATGACCGCTTCGATAC 1232
Qy 1251 GACACACCGCTTTTAAAGAGGTGCAATCTCTGGGGGGTTCGCGAGGAACATGCGGC 1310
Db 1233 GAACAGACGAGGTACAAAGTATCAAGGCGCTGGCTCCGCTTCGCAAGTCCAAACCGCGC 1292
Qy 1311 GATCCAGTACGCCACCAACCGAGGTTGGATCAAAATGATGTTTATATTTATCAACG 1370
Db 1293 TCTGCGCTAGGCTCGACGACCGAGCTTGGGTGACACGCGCTCTTACGTATATGACG 1352
Qy 1371 GAAATTTTCAACAGATGTCTGTTTGGTGGCCATCAATCGAAAACAGCAATCTCTCTATTTC 1430
Db 1353 CAAGTTCCGAGACCAACGTAGCTCTCGTTCGCGTCAACCGAGCTCGACGCTGCCTATCC 1412
Qy 1431 GATTTCCGGTTTGCAGACGCGCTTGCAAAATGCGAGCTATGCGGATATCTGTCAGGCT 1490

```



```

Db 1413 GATATCGGAGCGCTTACTGCTCTGCCAAACGGAACGATATACCGAGCTTCTCGCGCGCT 1472
QY GTTGGGGGGAACGGGATTCGGTTTCGAATGGAAGTGTGCTTCGTTCACGCTTCGGCC 1550
Db GCCTAATGGCAATCAATACGGTTAAGGCGGCACGGTCAGCAACTTTACACTTGCAGC 1532
QY TGGAGCGGTGTCGTTTGGCAGTACAGACATCCGTTTCAGCGCGGCAATTCGGATCGGT 1610
Db GGGCGGTACGGCAGTCTGGCAGTACAGCGACGGATCTTCGCGGATTTATCGCGAACGT 1592
QY TGTCTCAAAATATGGGATTCGGGTATGTGTACGATCGACGGGAAGGTTTGGGAC 1670
Db CGGCCGACTATGGCAGACCGCGCACACCATTCAGATCGACGACCGCGCTTCGGTAC 1652
QY GACCGAGGACCGTGACATTTGG-----CGGAGTGACAGCACTGTGAATTCCTG 1721
Db TACGAAGACAAAGTTACTTTTCGGTACGACGCGGTTACCGCGCGCAACATCGTAGCTG 1712
QY GACATCCATCGGATGGAAGTACGTTCCACATGCGCGCGGCTGACCGATGTGA 1781
Db TACGAAGACAAAGTTACTTTTCGGTACGACGCGGTTACCGCGCGCAACATCGTAGCTG 1712
QY GGAAGATACCGAATCAAGGTCGAAGTTCGAACGTCGCGCGCGCAACATCGCGCGTTAC 1772
QY AGTCACCGCGGTGGAGTTTCAGCAATCTGTATTCTTACAAATTT-----TTGAGTGG 1835
Db GGTAAACGAACGCGCGGCACTTACCGCGAGCGTTTCAACACTTTAAGTACTGACTGC 1832
QY AACCGACAGATCGGTTGTTTACTGTGAAAGTGGCTTCGACCACTCGGGGATAA 1895
Db CGATCAGGTACTGTCGCTTCAAGTCAAAGTCAAATGCGCACACGCG-----CTGGGACAAA 1889
QY GATTACTGACGCGGCAACATACCGAATTTGGGAATTTGGACACGGATACGAGCGGAGC 1955
Db CGTCTACCTGACCGGTACGTCGCGAGCTTGCACCTGG-----ACGCGCCCAA 1940
QY CGTTAAACATGCGCAAGGCGGCTGCTCGCGCCCAATTTACGAGTGGTTTATGATT 2015
Db CGCAATCGGTCCGATGTACCAACAGGTAGAAGCGAGCTATCGCACTTGTGACTTCGACGT 2000
QY CAGCGTTCAGCAGGAGAGAGATTCATCAATTCAGATTTCTCATCAAGCGTGGGATGAAC 2075
Db CAGGTTTCGGGCGCAACAGCGGCGTCAATTCAGATTTCTCAATCAAGTGAACCGGCTCGACGT 2060
QY GATTCAATGGGAGATGTTTGAACACCGTGGCCCAACTCCACCGGTTGCAACCGGTAA 2135
Db GACT---TGGGAAGGCGGCAACACCAACACTTCACTCGCTTCGAGCGCGGTTCGAC 2117
QY CATTACTGTTACGTGGCAAAAC 2157
Db CGTAACGGTTCGATTGGCAGAAC 2139

```

RESULT 12

```

E09507 LOCUS 2058 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding polypeptide which have cyclomaltodextrin
glucanotransferase activity from Bacillus macerans.
ACCESSION E09507
VERSION E09507.1 GI:22026134
KEYWORDS JP 1995163350-A/4.
SOURCE Paenibacillus macerans
ORGANISM Paenibacillus macerans
REFERENCE Sugimoto, T., Kubota, M. and Sakai, S.
AUTHORS DNA CODING POLYPEPTIDE HAVING CYCLOMALTODEXTRIN GLUCANOTRANSFERASE
TITLE ACTIVITY AND RECOMBINANT DNA CONTAINING THE DNA
JOURNAL Patent: JP 1995163350-A 4 27-JUN-1995;
HAYASHIBARA BIOCHEM LAB INC
COMMENT OS Bacillus macerans
PN JP 1995163350-A/4
PD 27-JUN-1995
PF 07-SEP-1994 JP 1994238623

```

```

PI SUGIMOTO TOSHIYUKI, KUBOTA MICHIO, SAKAI SHUZO PC
C12N15/00;C12N9/10;C12N15/00;C12R1:07;C12N9/10;C12R1:07;CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..2058
FT /organism="Bacillus macerans" FT
FT /strain="17A"
FT mat_peptide 1..2058
FT /note="Polypeptide which have
cyclomaltodextrin
glucanotransferase activity from Bacillus
macerans".
FEATURES
Location/Qualifiers
source 1..2058
/organism="Paenibacillus macerans"
/mol_type="genomic DNA"
/db_xref="taxon:44252"
BASE COUNT 517 a 635 c 524 g 382 t
ORIGIN
Query Match 16.18; Score 347.2; DB 6; Length 2058;
Best Local Similarity 52.2%; Pred. No. 1.1e-91;
Matches 1077; Conservative 0; Mismatches 913; Indels 72; Gaps 11;
QY 123 GGACGTGATTACAGATTATCATTCACCGGTTTACGATGGGACACGACGAAACAAA 182
Db 42 GGATACGGTTTACAGATTGTAACGACCGGTTTGTGACGG---CAATTCGCCACCA 98
QY 183 TCTGCGCAAAAGTTATGCGATTTCATCGACCGCAAAATCGAAGTGAATAATGTTTGGG 242
Db 99 CCGACCGGAGCAGCCTTCAGCAGCGATCATTTCCAACTGAACTGTATTTCGGGGCGA 158
QY 243 CGGGATCTGGAGGGGTTCGTCAAAAACCTTCCTTATCTTAAACAGCTGGGTGAACGAC 302
Db 159 CTGGCAGGGATCAGCAACAAATCAACGCGCTAICTGACGGAAATGGGATCACGC 218
QY 303 AATCTGGTTGTCCCGGTTTGGACAA-----TCTGGATACACTGGCGGACCGCA 353
Db 219 CCTCTGGATCTCGACGCGGTTGAGAAATCACCGCGCTCATCAATATTTCGGGGCTCAA 278
QY 354 TAACCGCGCTATCAGGATATCGACGCGGATTTTAAACAGATTGAGAACTTCGG 413
Db 279 CAATACGCTTACACCGGTTACTGGCTTCGAGCTTCAAGAACCAATGCCCGGTTGG 338
QY 414 GAATTGGACCACTTTGACACGTTTGTCTAATGANGCTCACCAAAACGGAATCAAGTGAT 473
Db 339 CAGCTTCACCGACTTCTCCAATTTGATCGCGCAGCGCATTCACAAATATCAAGGTAGT 398
QY 474 TGTGACTTTGTGCGCAATCATTCGACTCTCTTTTAAGGCAACAGATTCACCTTTGCGGA 533
Db 399 TATGACTTTGCACTCAATCACCAACCGGCTTCGAGTACGAGACCCCTCGTTGCGCGA 458
QY 534 AGCGCGCGCTCTACAAACAAATCGAACCTTATATGGCAATTTATTTGATGACGCAACAA 593
Db 459 GAACGGCGGCTCTACAAACAAACGGAACGCTGCTCGGCAAGTA---TAGCAACGATCCGC 515
QY 594 AGGTACTTCCACCAATATCGGACATCAGCAACTGGGACGACGCGTACGAGCGCAATG 653
Db 516 CGGCTGTTCACCAACAAATGGGCGCACCGATTCT-----CGACGAC 557
QY 654 GAATAACTTCAGGATCCAGCCGGTTTTCGCTTCCGATTTGTCGAGGAAATGGGAC 713
Db 558 TGAAGCGGTATCTACAGAACCTGTACGATCTCGCGGATATCAATCAGAACCAACAC 617
QY 714 GATTGCTCAATACCTGACCGATCGCGGCTTCAATTTGTGTAGCATGAGCGCGATGGTTT 773
Db 618 CATCGACTGCTATCTCAGGAATCGATCAGCTGTGGCTGAATCTCGAGTGCAGCGGAT 677
QY 774 GCGGATTGATCGGTGAAGCAATTTTAATTCGGGGTTCTCCAAATCGTTGCGCGATAAAT 833

```

D	b		1752	C	GATCAGGTCACCTCGCTTCAAGTCAAATGCCACCACGC---CCTGGGACAAA	180
Q	y		1896	G	ATTACTACGACGGGCAACATACCGGAATTGGGGAATTGGAGCACGGATPACGCGGAGC	1855
D	b		1809	C	GTCTACCTGACCGGTAACTCGCCGAGCTTTGGCAACTGG-----ACAGCGCCAA	1859
Q	y		1956	C	GTTCACATGGCGAAGGGCCCTGCTCGCGCCCAATTATCCGGATTGGTTTTATGTATT	2015
D	b		1860	C	GAATTCGGTCCGATGTACAACACAGGTAGAAGCCAGCTATCCGACTTGCTACTTCGACGT	1919
Q	y		2016	C	AGCGTTCCAGCGAAGAACGATTCAAATCAAGTTCTTCATCAACGCTGCGATFGGAAC	2075
D	b		1920	C	AGCGTTCCGCGCAACACGGCGCTGCAATTTCAAGTTTCATCAAAGTCAACGGCTGCACAGT	1979
Q	y		2076	G	ATTCATATGGGAGATGGTTTCGACACACAGCTGCGCACACTCCCAACGGGTGCACCGTTAA	2135
D	b		1980	G	ACT--TGGGAAGCGCGCAACCAACACACACTTCACCTCGCCTTCGAGCGCGTTGCGAC	2036
Q	y		2136	C	ATTACTGTTTACGTGGCAAAAC	2157
D	b		2037	C	GTAAACGGTTCGATTGGCAGAAC	2058
 RESULT 13 E12945						
LOCUS E12945 2058 bp DNA linear PAT 27-APR-1998						
DEFINITION DNA encoding cyclomaltodextrin glucanotransferase.						
ACCESSION E12945						
VERSION E12945.1 GI:3251776						
KEYWORDS JP 1997107968-A/2.						
SOURCE Paenibacillus macerans						
ORGANISM Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.						
REFERENCE Sugimoto, T., Kubota, M. and Sakai, S.						
AUTHORS METHOD FOR TRANSFERING SACCHARIDE						
TITLE Patent: JP 1997107968-A 2 28-APR-1997;						
JOURNAL HAYASHIBARA BIOCHEM LAB INC						
COMMENT OS Bacillus macerans						
PN JP 1997107968-A/2						
PD 28-APR-1997						
PF 03-DEC-1984 JP 1996265074						
PI SUGIMOTO TOSHIOYUKI KUBOTA MICHIO, SAKAI SHUZO PC						
C12N1/59, C12N1/21, C12N9/10, C12P19/18, (C12N1/21, C12R1:19), PC						
(C12N1/21, C12N9/10, C12R1:125), (C12N9/10, C12R1:07), (C12N9/10, C12R1:19), (C12N9/10,						
PC C12R1:125); CC C12R1:125);						
CC topologys: Double;						
CC topology: Linear;						
FH Key Location/Qualifiers						
FH source 1..2058						
FT /strain='17A',						
FT mat_peptide						
FT product='cyclomaltoedextrin						
FT glucanotransferase'.						
FEATURES						
source						
1..2058						
/organism='Paenibacillus macerans'						
/mol_type='genomic DNA'						
/db_xref='taxon:44252'						
BASE COUNT 517 a _335 c 524 g 382 t						
ORIGIN						
Query Match 16.1%; Score 347.2; DB 6; Length 2058;						
Best Local Similarity 52.2%; Pred. No. 1.1e-91;						
Matches 1077; Conservative 0; Mismatches 913; Indels 72; Gaps 11.						
Q	y		123	G	AGCGTATTACCATATCATTCAGTCGGTTCATCGATGGGACACGACGACACAA	182
D	b		42	G	ATACCGTTTACCATGATTGTAACCGACGGTTTGTGGACG---CAATTCCGCCAACCA	98

QY 183 TCCTGCCAATAAGTTATGAGCTTACGATCGGACCAATCGAAGTGGAAATGATTTGGGG 242
Db 99 CCCGACCGGAGCAGCTTTCAGACGCGATCATTTCCAACCTGAAGCTGATTTGGGGGGGA 158
QY 243 CGGSGATCTGGAGGGGGTTCGTCAAAACTTCTTAAACAGCTGGGGTAAACGAC 302
Db 159 CTGGCAGGGATCAAGAACAAATCAACGACGGCTATCTGACCGGAATGGGCAACCGC 218
QY 303 AATCTGGTGTCCCGGGTTTGGACAA-----TCTGGATACACTGCGGGCACCGA 353
Db 219 CCTCTGGATCTCGCAGCGGTTGAGACATCAACCGCGCTCATCAATATTCGGGGGTCAA 278
QY 354 TAACACGGGCTATCGAGTACTCGACGCGGATTTTAAACAGATTGAGGAACATTTGG 413
Db 279 CAATACAGCTTACCGAGTTACTGGCTCGGACTTCAAGAGACCAATCGCGGTTGG 338
QY 414 GAATTGGACCAATTTGACAGTTTGTCAATGATGCTACCAAAACCGAATCAAGTGTAT 473
Db 339 CAGCTTACCGACTTCTCCAATTTGATCGCGCAGCGATTCAACAATATCAAGGTAGT 398
QY 474 TGTGACTTTGTGCCAATCATTCGACTCCTTTTAAAGCAACGATTCACCTTTGCCGA 533
Db 399 TATGGACTTTGCACTTAATCAACCAACCGCGTTCGAGTACGGACCCCTCGTTGCCGA 458
QY 534 AGCGGGCGCCTCTCAACAATGAACTATATGCGCAATTTTATGACGCAACAAA 593
Db 459 GAACGGGGGCTCTACAAACGGAACGCTGCTCGGCAAGTA---TAGCAACGATACCGC 515
QY 594 AGGTACTTCCACONATATGGGACATCAGCAATGGGACCGGATCGAGCGCAATG 653
Db 516 CGGCGCTTCCACCAATGCGCGCACCGATTCT-----CGACGAC 557
QY 654 GAATACTTACGGATCCAGCGGTTCTCGCTTCGGATTTGTCGAGGAAATGGCAC 713
Db 558 TGAAGCGGTATCTAAGAACCTGTAGCATCTCGCGGATATCAATCAGAACAAACAC 617
QY 714 GATTGTCAATCTGACCGATGCGGGTTTCAATTTGATGATGACATGAGCGGATGGTT 773
Db 618 CATCGACTGTATCTCAAGGAATCGATCCAGCTGTGGTGAATCTCGGATCGACGGCAT 677
QY 774 CGGATGTAGCGGTGAGCATTTTAATTCGGGGTTCTCCAATTCGTTGGCCGATTAAC 833
Db 678 CCGCTTCGCGCGCTGAAGCATATGCTCAGGGCTGGCAGAGAGCTACGTCTCGTCTGAT 737
QY 834 GTACCAAA---AGAAAGACATTTTCTGTGGGGGAATGGTACGAGATGACCCCGGAAC 890
Db 738 CTACAGCAGCGCAATCCGGTGTTCACCTTCGTGATGTTCTCGGCCCCGACGAAT 797
QY 891 AGCCATCATCTGGAAAAGGTCCGGTACGCGCAACACAGCGGTGTCAATGTGCGATTT 950
Db 798 GACC-----CAGGACACATCAACTTCGCGAATCAGACGGGCATGCACCTGTGACTT 851
QY 951 TGATCTCAACAGCGGTGATTCGAATGTGTTGCGCACATTTAGCGAAACGATGTACATCT 1010
Db 852 TGGCTTTGCGCAGGAAATCCGTGAGTGTTCGCGCAACAGTCGAGGACGATGCCACT 911
QY 1011 TAACAATATGGTGAACCAACAGGGGAACGAGTACAAATACAAAGAAATTAATACAT 1070
Db 912 GAACTCGGTGATCTCAGCAACCGGCTCAGCTATAATTAATACAAACATGTTAGT 971
QY 1071 TATCGATAACCATGATCTCAGATTTCTTCGTTAAATTCGAAACGAGGGAATTTGCA 1130
Db 972 CATCGAACCATGATGATGACCGCTTCCAGAGCGGAGCGAGCATCTGCGCGACCGA 1031
QY 1131 CCAGCGCTGTCTTTCAATCTCAGCTTCGCGGGTACGCCCTCCATCTATTATGGAACCGA 1190
Db 1032 CGAGGCTTTGCGGTAACTGATCTTCGCGGGCTTCGGCAATCTACTACGATACAGA 1091
QY 1191 ACAATATATGGCAGGGCAATGACCCGTACAAACCGGGGATGATGCGGGTTGATAC 1250
Db 1092 GCAATATATGACCGGCAACCGGACCCGAAACCAACCGCGCATGATGACCGCTTCGATAC 1151
QY 1251 GACCAACACCGCCTTTAAAGAGGTGTCACTCTGCGGGGTTGCGCAGGAACATCGCGC 1310

Db 1152 GAACAAGACGCGCTACAAAGTGATCAAGGCGCTGGCTCCGCTTCGCAAGTCCAACCGGC 1211
QY 1311 GATCCAGTACGGCACCAACCCAGCGTTGGATCAACAATGATGTTTACATTTATGAACG 1370
Db 1212 TCTCGCTACCGGCTCGACGACCCAGCGTTGGGTGAACGACGACGCTTACGTAATGAACG 1271
QY 1371 GAAATTTTCAAGATGTCGTGTTGGTGGCCATCAATCGAAACACGCAATCCTCTATTC 1430
Db 1272 CAAAGTTGGGAAGCAAGTGTAGTCTCTGTTGCGTCAACCGCAGCTCGAGCTGCTATCC 1331
QY 1431 GATTTTCGGTTTCAGACGCGCTTCGCAAAATGGCAGCTATGCGGATTTATGTCAGGGCT 1490
Db 1332 GATATCGGGAGCGCTTACTGTCTCTCCCAACGAGAACTGTAACGACGTTCTCGCGGCT 1391
QY 1491 GTTGGGGGGAACCGGAATTCGCTTTCCAAATGGAAGTGTGCTTCTGTTCAACGCTTGCGCC 1550
Db 1392 GCTTAATGGCAATTTCAATTTACCGTTTAAACGGCGCAGCTCAGCAACTTTACATTTGACG 1451
QY 1551 TGGAGCGGTGTCTGTTTGGCAGTACAGCACATCCGCTTCAGCGCGCAAAATCGCATCGT 1610
Db 1452 GGGCGGTACGGCAGTCTGGCAGTACAGCAGCAGGAATCTCGCGATTTATCGGCAACGT 1511
QY 1611 TCTCTCCAAATATGGGAATTCGGGTAATGTGTCACGATCGACGGGAAAGGTTTGGGAC 1670
Db 1512 CGGCCCGGACTATGGCAAGCCCGCAACACCATCAAGATCGACGACGCGGCTTCGGTAC 1571
QY 1671 GACGAGGGAACGTCGACATTTGG-----CGGAGTACAGCGACTGTGAATCTGT 1721
Db 1572 TAGGAAGAACAAAGTTACTTTTCGTTACGACAGCGGTTTACCGCGCGCAACATCGTGAAGCTG 1631
QY 1722 GACATCCAAATCGGATTAAGTGTAGCTTCCCAACATGGCGCGCGGCTGACCGATGTGAA 1781
Db 1632 GGAAGATACGGAATCAAGTCAAGTTCGAAAGTTCGAAAGTGGCGCGCGCAACCGGCTTAC 1891
QY 1782 AGTCAACCGCGGTGGAGTTTCCAGCAATCTGTATTTTACAATATTT-----TTGAGTGG 1835
Db 1692 GGTAAAGAAACCGCGCGCACTACCGAGCGCAGCGTTCAACAATTTTAACTGACTGACTGC 1751
QY 1836 AACGACAGACATCGGTTGTTGTTTACTGTGAAAGTGCGCTCCGACCACTGGGGGATAA 1895
Db 1752 CGATCAGGTCTACTGTCTCGCTTCAAGTCAACAATGCCAACACCGC-----CTTGGGACAAA 1808
QY 1896 GATTTACCTGACGGCAACATACCGGAATTTGGGAATTTGGACGAGGATACGAGCGAGC 1955
Db 1809 CGTCTACCTGACCGGTAACGTCGCGAGCTTGCAACTGG-----ACAGCGCCAA 1859
QY 1956 CGTTTACATTCGCAAGGCGCTCTGCTCGCGCCAATTTATCCGATTTGTTTATGATTT 2015
Db 1860 CGCAATCGGTCGATGTACAAACGAGGTAGAAGCAGCTATCCGACTTGGTACTTCGACGT 1919
QY 2016 CAGCGTTCCAGAGGAAAGACGATTCATTCAGTTCTTCAACGCTGCGGATGGAAC 2075
Db 1920 CAGCGTTCCGCGCAACACGCGCTGCAATTCAGTTTCAATCAAGTGAACGGCTCGACGT 1979
QY 2076 GATTCATTTGGGAATGTTTCCGAAACAGTGGCCCAACTCCACGGGTGCAACGGTAA 2135
Db 1980 GACT---TGGGAAGCGCGCAACACCACTTCACTCGCTTTCGAGCGGTTGGAC 2036
QY 2136 CATTAAGTGTACGTGGCAAAAC 2157
Db 2037 CGTAACGCTCGATTGGCAGAAC 2058

RESULT 14

BACAMVRS
LOCUS 2996 bp DNA linear BCT 26-APR-1993
DEFINITION Bacillus sp. raw-starch-digesting amylase gene, complete cds.
ACCESSION M33302
VERSION M33302.1 GI:142508
KEYWORDS amylase.
SOURCE Bacillus sp.
ORGANISM Bacillus sp.

REFERENCE

Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

1 (bases 1 to 2996)

Ikhor, P., Tsukagoshi, N. and Udaka, S.

NUCLEOTIDE SEQUENCE OF THE STARCH-DIGESTING AMYLASE GENE FROM
BACILLUS SP. B1018 AND ITS STRONG HOMOLOGY TO THE CYCLODEXTRIN
GLUCANOTRANSFERASE GENES

JOURNAL Biochem. Biophys. Res. Commun. 166 (2), 630-636 (1990)

MEDLINE 90147765

PUBMED 1689153

Original source text: Bacillus sp. (strain B1018) DNA, clones

pUC19-A[1,2,3].

FEATURES

source

Location/Qualifiers

1..2996

/organism="Bacillus sp."

/mol_type="genomic DNA"

/db_xref="taxon:1409"

313..2454

/note="raw-starch-digesting amylase precursor"

/codon_start=1

/transl_table=11

/protein_id="AAA22239.1"

/db_xref="gi:142509"

/translation="MKELKMTAPSLGSLAFLSPQAAPDTSVNKNFSDVI
YQFTDRSDGNPANTGAPDGTCTNLRLCYGSDAGLKNKINDYLTGMYTAIW
ISQFVENIISGVNNTAYRGLYWARDFKINPAITGTFADFNLIAAHAKNIV
IDFAPNHTSPASSQDFSAENGRLYDNGTLGLGYTNDTONLFHHNGTDFSTENGII
KNLYDLADLNHNNTSDVLYLKDIAIKWLDLGDGIRMDAVKHMPFGWQSFMAVNNY
KPVTFGEWFLGVEVEPNEKFNESGMSLLDFRPAQKRVPRFNDTNDMYGLKML
EGSAADYADDDQVTFIDNHMERFHASNARKLEQALFTLILARVEAIYVGEQY
MSGTDPDNRARI PSFTSTTAYVIOKLAPLRKSNPAIYAVGTORRNNNDVLYER
KFGSNVAVVNNELNAPASISGLVLSLPGSYNDVNLGLNGNTLVGSGAASNP
LAAGTAVNQVTAIRPTPTIGHVGNMAKPGVTIIDRGFGSGSKTVTFGTATVSGA
NITSWEDTQIKVPIAVAGGIYIKVANAAGTASVNDYDNEFVLSGQVDSRVVNNAT
TALGQNLVLTGVSLEGNWDPFAKAIQENYQVYQYPNWYDVSVPAGKTIEFKPLKK
QGSTVWEGSGSNHTFAPSSTGTATINWQP"

sig_peptide /note="raw-starch-digesting amylase signal peptide"

mat_peptide /note="raw-starch-digesting amylase"

/product="raw-starch-digesting amylase"

BASE COUNT 764 a 852 c 774 g 606 t

ORIGIN

Query Match 16.0%; Score 345; DB 1; Length 2996;
Best Local Similarity 51.9%; Pred No. 5.6e-91;
Matches 1075; Conservative 0; Mismatches 920; Indels 78; Gaps 10;

QY 111 AAGCGTCAAGGGAGCGTGATTACAGATATACATGACCGGTTTACGATGGGACAC 170

DB 423 AAATTCAGCACCGACGTCATCTACAAATTTTCACCGACAGGTTTTCGAGCGCAATCC 482

QY 171 GACGAACAACATCTCGCCAAAGTTATGACTTTACGATCCGACCAATCGAAGTGGAA 230

DB 483 CGCAACAATCCAGCGGC-----GCGCGTTTGACGGAACCTCGACAACTCCG 533

QY 231 AATGTATTGGGCGGGGATCTCGAGGGGTTGTCFAAAAAC-----TCCTATCTTAA 284

DB 534 GCTGTATTGGGCGGGGACTGCGAGGTCATCATCAACAAATCAACGACGCTACCTGAC 593

QY 285 ACAGCTGGGCTAACACATCTGTTGTCGCCGGTTTGGACATCTGGATACCT--- 341

DB 594 CGGATGGGGGTTTACCGGCATCTGATCTCCCGCGGTGGAACATCTACAGATCAT 653

QY 342 -----GCGCGGACCGCATAAACGCGGTATACAGGATCTGAGCGCGGATTTTAAACA 395

DB 654 CAATTAATTCGGCGTCACACACACGCGCTATCACGGCTACTCGGCGCGGACTTCAGAA 713

QY 396 GATTGAGGAACATTTGGGAATTGACACATTTGACACGTTGGTGCATATGATGCTACCA 455

DB 714 GACGAATCCGCGCTACGCGACGATTCGCGACTTCCAGACCTGATCGCGCGCGCATGC 773

QY 456 AAACCGAATCAAGTGATGTTGACATTTTGCCCAATCATCTGACATCTCTTTTAAGGCAAA 515

DB 774 CAAAACATCAAGTCTATATGACATTTGCGCCCGAACCATATACGTCCCGCGCTCTCGGA 833

QY 516 CGATTCCACCTTTGCGGAAGCGGCCCTCTACAACATGGAACCTATATGGCAATTA 575

DB 834 CAGCGTTTCCTTTTCCGGAAACGGCGCTGTAGATACGACGCTGCTCGGGGATA 893

QY 576 TTTTGATGACGACCAACAAAAGGTACTTCCACCATAAATGGGACATCAGCAACTGGACGA 635

DB 894 CACGAACGATACGCAGAACCTGTTCACGACATAAC-----GGCGG 932

QY 636 CGGTACGAGCGCAATGGAATAAATTTCAGGATCCAGCGGTTTCTGCTTGCCTGCGGATTT 695

DB 933 CAGGACTTTTCCAGACCGAAACGGCATCTCAAAAACCTGTAGATCTCGCGGACCT 992

QY 696 GTCCAGGAATAATGGCACGATTGCTCAATACCTACCGATCGCGGCTTCAATTGGTAGC 755

DB 993 GAACCATAAACAACGACGCTCGGACGTCTACTTGAAGGACGCGATCAAAAATGTGGCTGA 1052

QY 756 ACATGAGACGGATGTTGGGATTAATGCGGTGAAGCATTTTAATTCGGGTTTCTCAA 815

DB 1053 TCTCGGCATCGAAGGCATCCGATGATGCGGTGAAGCATATGCGGTTTCGGCTGGCGAA 1112

QY 816 ATCGTTGGCCGATAAATCTGTACAAAAGAAAGACATTTCTCGTGGGGAATGTTACGG 875

DB 1113 GAGCTTTATGCTGCCGTCAACATTAAGCCGGTCTTACCTTCGCGGATGTTCTCT 1172

QY 876 AGATGACCCCGGAACAGCAATCATCTGAAAAGGTCCGGTACGCGACACACAGCGGTGT 935

DB 1173 GGGCGTAAACGAAGTAGGC-----CGGAAACCATTAAGTTTGCCAAACGAATCCGGCAT 1226

QY 936 CAATGTCTGATTTGATCTCAACACGGTCATTCGAAATGTGTTCGGCACATTTACGCA 995

DB 1227 GAGCTTGTTGATTTCCGTTTGGCCAAAAGTGGCGAGGTGTTCCGGGACACACCGA 1286

QY 996 AAGCATGTACGATCTTAACAATATGTTGAAACCAACCGGGAACAGTACAATAACAAGA 1055

DB 1287 CAATATGTACGGCTGAAGCGCATCTGAGGGGTCCGACGCGATTACGCCAGGTGA 1346

QY 1056 AATCTAATCATATTTATGATAACCATGATATCTCAAGATTTCTTGGTAAATTCGAA 1115

DB 1347 TGACAGGTGACGTTTCATGACAAACATGACATGAGCGGTTTCCAGCAAGCAATGCAA 1406

QY 1116 CAAGCGAATTTGACACAGCGCTTGTCTTTCAATCTCACTTCGCGGGGTACGCCCTCAT 1175

DB 1407 CCGCGGAAGCTGAGCAACGCTGCGGTTCAACGCTGATCTCGCGCGCTCCCGCCAT 1466

QY 1176 CTATTATGGAACCGCAACAATACATGCGAGCGGCAATGACCCGTACAAACCGGGGATGAT 1235

DB 1467 TTATTACCGCACCGGACGATACATGTCGGGTGGGACCGATCCGGACAAACCGGGCGCGAT 1526

QY 1236 GCGCGGTTTGATACGACAAACACCGCTTTTAAAGAGGTGTCACTCTGCGGGGTTGCG 1295

DB 1527 CCGTCTCTCCAGCTCGACGACGCGCTATCAAGTCATTTCAAAGCTGCGCGCGCTGG 1586

QY 1296 CAGGAACAATGGCGGATTCAGTACGACACACACCGCGGTGGATCAACAATGATGT 1355

DB 1587 CAAGTCAACCGCGCATCTGCTACGATTCGACGAGCGCGTGGATCAACACGACGT 1646

QY 1356 TTACATTTATGAACGAAAATTTTCAAAGATGTCGTGTGTGSCCATCAATCGAAACAC 1415

DB 1647 GCTCATTTATGAGCGCAATTCGGACGCTTGTCACTCTCCCGCAAGGACGCTACAATGA 1706

QY 1416 GCAATCTCTTATTCGATTTCCGGTTTGACACGCGCTTCCCAAATGGCAGCTATGGGA 1475

DB 1707 GAACGCGCGCTCCATTTCCGGACTTGTCACTCTCCCGCAAGGACGCTACAATGA 1766

QY 1476 TTAATCTGACGGCTGTGGGGGGGAAACGATTTCCGTTTCAATGGAAGTGTGCTTC 1535

DB 1767 CGTCTTGGCGGCTCTGAACCGCAACGTTAACGCTAGGCTCCGCGGAGCGCGCTC 1826

QY 1536 G---TTCAAGCTTCGCGCTGAGCGCTGCTGTTTGGCAGTACAGCACATCGCTTCAGC 1592

DB 1827 CAATTTTACGCTTTCGCGCGCGGACGCGGCTGTGGAGTGTACCGCGCAACCGGAC 1886

```

/strain="B1018"
/db_xref="taxon:417"
/note="10 bp upstream of RsaI site."
101..106
-35 signal
-10 signal
125..130
RBS
302..306
note="ribosome binding site"
313..2454
/EC number="3.2.1.1"
/function="alpha amylase"
125..130
/codon_start=1
/transl_table=11
/product="raw-starch-digesting amylase precursor"
/protein_id="BAA4140.1"
/db_xref="GI:216325"
/translation="MKPKLMTAFSLGLSLAFSPAAQADTSVSNKQNFSTDIV
YQIFLRFQDGNPNPTGAAFDGCTNLRLYCGDMQGIINKINDGLYLGWGYTALE
ISAPVENIYSILNYSGVNTAYHYGWARDFKTNPTAGTADFQNLIAAAHAKNIKVI
IDFAPNHTSPASSDQPGFPAENGPLYDNGTLTGGYTNQTLPHHGSGTDFPSTENGILY
KULYDLADLNHNSTSDVYIKDALIKWLDGLDGRMDVAKGMPFGWQKSPFAVANNY
KSPFTGEWFLGVNVEPENEHKFANESGMSLLDFPQAKVRQVFRDNTDNNMGLKAVL
EGSAADYQDDQVTFIDNHMERFTHANANRLEQLAFTLILARVPAIYYGTEGY
MSGGTDFPNRAIRFSFTSTAYTQIKLAELRKSNPAIAYSTQERMINDVLIYER
KFSNVAVAVNRNLNPAISIGLTVLPQGSYNDVLGGGLNGTLITVSGGTAASNP
LAAGTAWQNYTAAPTPTIGHVPMWAKPGVTITIDGRCPSSKGTIVYRCGTAVSA
NITSWEDTQIKYKIPAVAGIYNIKVANAAGTAGNVYDNEVLSDQSVSRFVNNAT
TAGLQNYLTGNVELNGDPAKAGTGNQVTVQIPFNWYIDVSPAGKTIEFPLFKK
CGSTVWEGGSGNHFTFAPSSGATINWVWQP"
313..393
/note="raw-starch-digesting amylase signal peptide"
394..2451
/product="raw-starch-digesting amylase"
2567..2580
/rpt_type=inverted
2585..2598
/rpt_type=inverted
COUNT 764 a 852 c 774 g 606 t

```

Query Match	16.0%;	Score 345;	DB 1;	Length 2996;
Best Local Similarity	51.9%;	Pred. No. 5.6e-91;		
Matches 1075;	Conservative	0;	Mismatches 920;	Indels 78; Gaps 10;
QY	111	AAGCGTCAAAGGGGACGTGATTTACAGAGTATATCATTTGACGCGTTTACGATGGGACAC	170	
Db	423	AAATTTACGACCGGACGTCATCTATCAAAATTTTACGACAGGTTTTCGGACGGCAATCC	482	
QY	171	GACGACAAACAATCTGCGCAAAAGTTATGGAATTTTACGATCCGACCAAAATCGAATGGAA	230	
Db	483	CGCCACAATCCGACCGGC-----CGGCGGTTTGAGGAACCTGCACGAACCTCCG	533	
QY	231	AAATGATTTGGGCGGGGATCTGAGGGGTTGCTCAAAACT-----TCCTTATCTTAA	284	
Db	534	GCTGTATTTGGCGGGGACCTGGCAGGGCATCATCACAAAATCAACGACGGTTACTCTGAC	593	
QY	285	ACAGCTGGCGGTAAACGAACATCTGTGTGTCGCCGGTTTTTGGACAATCTGGAATACACT---	341	
Db	594	CGGGAATGGCGTTACCGGCATCTGGATCTCCACGCCGTCGAAAAATCTACAGCATCAT	653	
QY	342	-----GGCGGGCACCGANAAACCGGGCTATACGGATATCTGGACGCGGGATTTTAAACA	395	
Db	654	CAATATTTCCGCGCTCAACACACGCGCTTATCAGGCTACTGGGCTACTGGGCGCGGACCTTCAAGAA	713	
QY	396	GATTGAGGAACATTTTCGGGAATTTGGACACACATTTGACACGTTGGTCAATGATGCTCACCA	455	
Db	714	GACGAATCCGGCTACGGACGAGTTCCGAGCTTCCAGAACCTGATCGCGCGCGGCGATGC	773	
QY	456	AAACGGGAATCAAGGTGATGTTCTGACTTTTGTGCCCAATCATTTTCGACTTCCTTTTAAAGGAAA	515	
Db	774	CAAAACATCAAGTCAATATCGACTTCGCCCGGACCAATACGTCGCCCGCGCTCGTCCGA	833	
QY	516	CGATTCACCTTTGCGAAGCGGCGCGCTCTTCAACAATGGAACCTATATGGGCAATTA	575	

Db 834 CCAGCCTTCCTTTTGGGAAAAACCGCGCTGTACGATAACGCGACGCTGCTCGGGGATA 893
QY 576 TTTTGTATGACGCAACAAAGAGGTACTTCCACCAATATGGGACATCAGCAACTGGGACGA 635
Db 894 CAGCAACGATACGAGAACCTGTTCACCAATAAC-----GGCGG 932
QY 636 CCGGTACGAGGGCAATGGAAAAATTTACGGATCCAGCCGGTTTCTCGTTCGCCGATTT 695
Db 933 CACGGACTTTTCCAGCCGGAACCGCATCTACAAAACCTGTACGATCTCCGCCACT 992
QY 696 GTGCGAGGAATGCGACGATGTCTCAATACCTGACCGATGCGGGGTTCAATGTGTAGC 755
Db 993 GAACCAATAACAACAGCAGCTCGGACGTCTACTTTGAAGGACGCGATCAAAATGTGGCTGA 1052
QY 756 ACATGAGCGGATGTTTGGGATTTGATCGGTGAAGCAATTTAATTCGGGGTTCTCCAA 815
Db 1053 TCTCGGCATCGAGGCATCCGCTGATGCGGTGAGCAATATGCGGTTCGGCTGGCAGAA 1112
QY 816 ATCGTTGGCGGATAAAGTGTACCAAAAGAAAGACATTTTCTGTGTGGGGAATGGTACGG 875
Db 1113 GAGCTTTATGGCTGCGCTCAACAACTATAAGCCGGTCTTTTACCTTCGGCGAATGGTTCT 1172
QY 876 AGATGACCCCGGAAACAGCCATATCTGGAAGAGGTCCGGTACGCGCAACACAGCGGTGT 935
Db 1173 GGGCGTAACGAAGTAGGC-----CCGGAACCCATAAGTTTGGCCACGATCCCGGAT 1226
QY 936 CAATGTGCTGGATTTTGATCTCAACACGGTGATTCGAATGTGTTCGGCACATTTACGGA 995
Db 1227 GAGCGTGTGATTTCCGTTTGGCCAAAGGTGCGGAGGTTCCTCGGACCAACACCGA 1286
QY 996 AACGATGTAGATCTTTACAAATATGTGTAACCAACCGGGAACGATGACAAATACAAAGA 1055
Db 1287 CAATATGTACGCGCTGAAGGGGATGCTGGAGGGCTCCGACGCCGATTAAGCCAGAGTGA 1346
QY 1056 AATCTAATCACAATTTATCGATAACCATGATATGTCAAGATTTCTTCGGTAATTCGAA 1115
Db 1347 TGACAGGTGACGTTTCATCGACACCAATGATGATGAGAGCGTTTCACGCAAGCAATGCAA 1406
QY 1116 CAAGCGGAATTTGCAACAGGCGGTGTGTTTCAATCTCAATTCGCGGGGTACGCCCTCCAT 1175
Db 1407 CCGCGGAGCTGGAGCAAGCGTGGCGTTTACGCTGATCTCTCGCGCGCGTCCCGCCAT 1466
QY 1176 CTATATGGAACCAACATACATGAGCGGCGCATGACCGGTACACACCGGGGATGAT 1235
Db 1467 TTATACGCGCACCGAGCAGTACATGTGCGGTGGGACCGATCCGGACAAACCGGGCGCGGAT 1526
QY 1236 GCGGCGTTTGATACGACAAACACCGCTTTTAAAGAGGTGTCAACTCTGCGGGGTTGGG 1295
Db 1527 CCTTCTCTTCCACGTTCGACGACCGCTATCAAGTCAATCAAAAGCTGGCGCGCTGGG 1586
QY 1296 CAGGAACAATGCGGCGATCCAGTACGGCAACACCGAGCGTTGATCAACATGATGT 1355
Db 1587 CAAGTCCAAACCCGGCCATCGCCTACGGATCGACGAGGAGCGCTGGATCAACACAGCST 1646
QY 1356 TTACATTTATGAACGGAATTTTTCACGATGTCGTTTGGTGGCCATCAATCGAAACAC 1415
Db 1647 GCTCATTTATGACGGAATTCGGGACGACGTTGCGGTGCTGTCGCTCAACCGCAATTT 1706
QY 1416 GCAATCTCTCTATTCGATTTTCGAGAGCGGCTTGGCAAAATGGCAGCTATGCGGA 1475
Db 1707 GAAACGCGCGGCTTCATTTTGGGACTTGTCACTTCCCTGCGGGAAGCAGCTACAAATGA 1766
QY 1476 TTATCTGTGAGGCTGTTGGGGGGAACGGGATTCGGTTTCGAATGAAGTGTGCTTC 1535
Db 1767 GGTCTTTGGGCGCTTCTGAACGGCAACACGTTAACGGTAGGCTCCGGCGAGCGGCTC 1826
QY 1536 G---TTACCGCTTGGCGCTGGAGCGCTGCTGTTTGGCAGTACAGCATCCGCTTCAGC 1592
Db 1827 CAATTTACCGCTTGGCGCGCGGCAACGGCGTGTGGCAGTACCGCGGCAACCGGAC 1886
QY 1593 GCGCAAAATCGGATCGGTTGCTTCCAAATATAGGGGATTTCCGGGTAAATGTGTACAGATCGA 1652
Db 1887 GCGGACCATCGGCGATGTCGGCGCGATGATGGCCAAAGCCGGGCGTGACGATCAGATCGA 1946

QY 1653 CCGGAAAGGTTTTTGGGACCAACCGGAGAACCGTGACATTTTGGCGAGTGACAGCGACTG- 1711
Db 1947 CGGCGCGGCTTCGGCTCTAGCAAAAGGCACCGTACTTTCGGCACGACGCGGTGAGCGG 2006
QY 1712 -----TGAATCTCTGGACATCCATCGGATTTGAAGTGTAGCTTCCCAACATGCGCGC 1763
Db 2007 CGCAACATCATCGTCTTTGGGAGACACGCGAGATCAAAAGTGAATAATTCGGCGCGTGCAGG 2066
QY 1764 CCGGCTGACCGCATGTGAAAGTC---ACGCGGGTGGAGTTTCCAGCAATCTGTATTCTTA 1820
Db 2067 CGGATCTCAACATTAAGTTCGAACCCCGCGGACCGGACGCAAGCTGTAGCAAA 2126
QY 1821 CAATATTTTGAAGTGAACGCGACATFCGGTTGTGTTTACTGTGAAAAGTGCCTCCGAC 1880
Db 2127 CTTGAGAGGTATTGTCCGGAGACCGAGTCCAGCTCCGCTTCGTGTCAACAACGCGACAAC 2186
QY 1881 CAACCTGGGGGATTAAGATTTTACCTGACGGGCAACATACCGGATTTGGGAATTTGAGCAC 1940
Db 2187 GCGCTTGGGCAAAATCTTACCTGACGGGCAATGTACGAGCTGGGAATGGGACCC 2246
QY 1941 GGATACGAGCGAGCGCTTAACAATCGCAAGGGCCCTGCTCGCGCCCAATTTATCCGGA 2000
Db 2247 G-----GCAAAAGCGATCGGCGGATGTACACAGGTGTTTACCAATATCCGAA 2297
QY 2001 TTGTTTTATGATTCAGCGTTCCAGCAGGAAGACGATTCATTCAGTTCTTTCATCAA 2060
Db 2298 CTGTTATTCAGCTCAGCGTTCCGGCCCGCAAAACGATCGAGTTCAGTTTTCAGAAAA 2357
QY 2061 GCGTGGGATGGAACGATTCATATGGGAGATGTTTCGAACCGCTGGCCCAACATCCAC 2120
Db 2358 AC---AAGGCTCCACCGTACGTTGGGAAGCGGACGACACACCTTCACCGCGCGCTC 2414
QY 2121 GGTGCAACCGGTAAACATTTACTGTTACGTGGCA 2153
Db 2415 CAGCGCACCGCGACCATTTAACGTGAATTTGGCA 2447

Search completed: September 17, 2003, 09:58:57
Job time : 5284 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: September 17, 2003, 06:20:16 ; Search time 420 Seconds
 (without alignments)
 13882.824 Million cell updates/sec

Title: US-10-069-908-1
 Perfect score: 2160
 Sequence: 1 atgaaaagaagaacgcttc.....ctgtacgtggcaaaactag 2160

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
 Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : N Geneseq_19Jun03.*
 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
 25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2160	100.0	2160	20 AAX87877	Bacillus sp. malto
2	2160	100.0	2160	20 AAZ10579	DNA encoding a mal
3	2160	100.0	2160	21 AA27399	Bacillus maltogeni
4	2160	100.0	2160	22 AAF62389	Bacillus maltogeni
5	2160	100.0	2160	22 AAF77165	Bacillus maltogeni
6	576	26.7	576	19 AAV28880	Nucleotide sequenc
7	576	26.7	576	20 AAX23988	B. stearothermophi
8	348.8	16.1	2058	7 AAN60739	Sequence encoding

9	343.4	15.9	2596	12	AAQ12366	Gene encoding enzy
10	328.6	15.2	2414	14	AAQ50345	Wild type cyclomal
11	327	15.1	327	19	AAV28879	Nucleotide sequenc
12	327	15.1	327	20	AAV23987	B. stearothermophi
13	326.2	15.1	2565	11	AAQ01810	Sequence encoding
14	324.6	15.0	2659	12	AAQ10164	Cyclomaltodextrin
15	321.8	14.9	2040	18	AAV85105	Cyclomaltodextrin
16	318.2	14.7	2040	17	AAV44121	Wild type Cgrase c
17	317	14.7	2040	7	AAN60737	Sequence encoding
18	304	14.1	2686	12	AAQ10163	Cyclomaltodextrin
19	301.4	14.0	2687	11	AAQ01792	Sequence encoding
20	296	13.7	2061	15	AAQ77669	Variant cyclomalto
21	281	13.0	392	21	AAQ77772	Fusarium venenatum
22	265.6	12.3	2340	21	AAZ56902	Cyclodextrin forma
23	263.2	12.2	4785	14	AAQ48794	B.obbensis Ala(188
24	263.2	12.2	4785	14	AAQ48795	B.obbensis Ser(188
25	262.8	12.2	4785	14	AAQ48788	B.obbensis mutant
26	262.6	12.2	4785	14	AAQ48797	B.obbensis Trp(188
27	261.4	12.1	4785	14	AAQ48796	B.obbensis Ser(188
28	255.4	11.8	4785	12	AAQ14072	Cyclodextrin gluca
29	254	11.8	4785	12	AAQ14942	Cyclomaltodextrin
30	245.6	11.4	1580	19	AAV19363	Plasmid pTN603 enc
31	243.8	11.3	3442	7	AAN60705	B. agaxaderens CG
32	236.4	10.9	2139	24	AAV88713	Bacillus cyclodext
33	236.4	10.9	2139	24	ABN85628	B. agaxaderans c
34	234.6	10.9	2631	24	ABK10103	Bacillus clarkii c
35	233.2	10.8	2530	25	ABZ82056	B. agaxaderans c
36	231.4	10.7	2175	24	ABA99462	Actinoplanes sp DN
37	231.4	10.7	17955	19	AAV56642	Actinoplanes sp. a
38	231.4	10.7	38064	24	ABA99469	Actinoplanes sp SE
39	229.8	10.6	2582	18	AAV73117	Actinoplanes sp. a
40	211.4	9.8	2133	24	ABN66655	Streptococcus poly
41	208.2	9.6	2100	12	AAQ13985	Gamma-cyclodextrin
42	154	7.1	1155	19	AAV19370	Cyclomaltodextrin
43	110.2	5.1	2112	22	AAV74682	P. furiosus thermo
44	110.2	5.1	2169	22	AAV74678	P. furiosus thermo
45	110.2	5.1	3519	22	AAV74681	P. furiosus thermo

ALIGNMENTS

RESULT 1
 AAX87877
 ID AAX87877 standard; DNA; 2160 BP.
 XX
 AC AAX87877;
 XX
 DT 22-NOV-1999 (first entry)
 XX
 DE Bacillus sp. maltogenic alpha-amylase Novamyl gene.
 XX
 KW Alpha-amylase; maltogenic; cyclodextrin glucanotransferase;
 KW Novamyl; Cgrase; enzyme engineering; baking; anti-staling; ss.
 XX
 OS Bacillus sp.
 XX
 FH Key Location/Qualifiers
 FT sig_peptide 1..99
 FT /*tag= a
 FT mat_peptide 100..2157
 FT /*tag= b
 FT
 XX
 XX WO9943793-A1.
 PD 02-SEP-1999.
 XX
 XX 26-FEB-1999; 99WO-DK00087.
 XX
 PR 12-MAR-1998; 9AUS-0077795.
 PR 27-FEB-1998; 98DK-0000269.
 PR 27-FEB-1998; 98DK-0000273.
 PR 11-MAR-1998; 98US-0077509.

QY 1561 TCTGTTGGCAGTACAGCACATCCGCTTCAGCGCCGCAATCGATCGGTCCTCCAAAT 1620
 Db |||||
 1561 TCTGTTGGCAGTACAGCACATCCGCTTCAGCGCCGCAATCGATCGGTCCTCCAAAT 1620
 QY 1621 ATGGGGATTCGGGTAAATGTGTCTACGATCAGCGGGAAGGTTTTGGACGACGAGGGA 1680
 Db |||||
 1621 ATGGGGATTCGGGTAAATGTGTCTACGATCAGCGGGAAGGTTTTGGACGACGAGGGA 1680
 QY 1681 ACCGTGACATTTGGCGAGTGCACGAGCTGTGAATCTCTGGACATCCCAATCGGATTGAA 1740
 Db |||||
 1681 ACCGTGACATTTGGCGAGTGCACGAGCTGTGAATCTCTGGACATCCCAATCGGATTGAA 1740
 QY 1741 GTGTACGTTCCCAACATGSCCGCGGCTGCACGATGTGAATGTCACTCCGCGGTGGAGTT 1800
 Db |||||
 1741 GTGTACGTTCCCAACATGSCCGCGGCTGCACGATGTGAATGTCACTCCGCGGTGGAGTT 1800
 QY 1801 TCCAGCAATCTGATTCTTACAAATATTTTGTAGTGAACGACGACATCGGTTGTGTTACT 1860
 Db |||||
 1801 TCCAGCAATCTGATTCTTACAAATATTTTGTAGTGAACGACGACATCGGTTGTGTTACT 1860
 QY 1861 GTGAAAGTGGGCTCGACCAACTGCGGGGATTAAGATTTTACCTGACGGGCAACATACCG 1920
 Db |||||
 1861 GTGAAAGTGGGCTCGACCAACTGCGGGGATTAAGATTTTACCTGACGGGCAACATACCG 1920
 QY 1921 GAATTGGGAATTTGGAGCAGCATACGAGCGGACCGCTTAACATGCGCAAGGCGCCCTG 1980
 Db |||||
 1921 GAATTGGGAATTTGGAGCAGCATACGAGCGGACCGCTTAACATGCGCAAGGCGCCCTG 1980
 QY 1981 CTCGGCCCAATTATCGGATTTGTTTTATGTATTACGCTTCCAGAGGAAAGACGATT 2040
 Db |||||
 1981 CTCGGCCCAATTATCGGATTTGTTTTATGTATTACGCTTCCAGAGGAAAGACGATT 2040
 QY 2041 CAATTCAGATTTCTATCAAGCTGCGGATGCAAGATTCATTTGAGAGATGCTTCGAC 2100
 Db |||||
 2041 CAATTCAGATTTCTATCAAGCTGCGGATGCAAGATTCATTTGAGAGATGCTTCGAC 2100
 QY 2101 CAGTGGCCCAACTCCACGGGTGCAACCGGTAACTACTGTTACGTGCAAAACTAG 2160
 Db |||||
 2101 CAGTGGCCCAACTCCACGGGTGCAACCGGTAACTACTGTTACGTGCAAAACTAG 2160

RESULT 2

AAZ10579
 ID AAZ10579 standard; DNA; 2160 BP.
 XX AAZ10579;
 AC AAZ10579;
 DT 17-NOV-1999 (first entry)
 XX DNA encoding a maltogenic alpha amylase.
 DE
 DE Maltogenic alpha-amylase; variant; dough additive; bread staling;
 XX search retrogradation; organoleptic property; detergent; sweetener;
 KW linear oligosaccharide production; ethanol; starch; textile desizing; ss.
 XX
 OS Bacillus sp.
 XX
 FH Key Location/Qualifiers
 CDS 1..2160
 FT /*tag= a
 FT sig_peptide 1..99
 FT /*tag= b
 FT mat_peptide 100..2157
 FT /*tag= c
 FT
 XX WO9943794-A1.
 FN
 PN
 PD 02-SEP-1999.
 XX
 PF 26-FEB-1999; 99WO-DK00088.
 XX
 PR 27-FEB-1998; 98DK-0000269.

PR 12-MAR-1998; 98US-0077795.
 XX (NOVO) NOVO-NORDISK AS.
 PA Cherry JR, Svendsen A, Andersen C, Beier L, Frandsen TP;
 XX WPI; 1999-550868/46.
 DR P-PSDB; AAY30621.
 XX
 PT New variants of maltogenic alpha-amylase, particularly used as dough
 PT additives to retard staling
 XX
 PS Claim 1; Page 133-137; 145pp; English.
 XX
 CC The present sequence encodes a maltogenic alpha-amylase. The
 CC specification describes variants of maltogenic alpha-amylases,
 CC which are at least 70% identical to the encoded protein. The
 CC variants have altered pH-dependent activity, improved stability
 CC or higher specific amylase activity than the original amylase.
 CC The variants are used as dough additives to reduce retrogradation
 CC of starch and/or staling of bread. In bread they should result in
 CC increased volume and improved organoleptic properties (flavor,
 CC mouthfeel, palatability, aroma and crust color). Other uses
 CC are in detergents; production of linear oligosaccharides;
 CC production of sweeteners or ethanol from starch; and textile
 CC desizing.
 XX
 SQ Sequence 2160 BP; 571 A; 499 C; 572 G; 518 T; 0 other;
 Query Match 100.0%; Score 2160; DB 20; Length 2160;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAAAAGAAAACGGTTTCTTTATTTGTGGACTGATGCTCTCATCGTCTTCTGTTTC 60
 Db |||||
 1 ATGAAAAGAAAACGGTTTCTTTATTTGTGGACTGATGCTCTCATCGTCTTCTGTTTC 60
 QY 61 AGCGGTCTCTTCCTCGTACAAATCCAAAGCCGCTGAGCCAGCAGTTCGCGAAGCGTCAAA 120
 Db |||||
 61 AGCGGTCTCTTCCTCGTACAAATCCAAAGCCGCTGAGCCAGCAGTTCGCGAAGCGTCAAA 120
 QY 121 GGGGACGTGATTTACCAAGATTCATTGACCGGTTTTACGATGGGGACACGACCAACAAC 180
 Db |||||
 121 GGGGACGTGATTTACCAAGATTCATTGACCGGTTTTACGATGGGGACACGACCAACAAC 180
 QY 181 AATCTGCCAAAAGTTATGGACTTTTACGATCCGACCAATTCGAAGTGAATAATGTTTGG 240
 Db |||||
 181 AATCTGCCAAAAGTTATGGACTTTTACGATCCGACCAATTCGAAGTGAATAATGTTTGG 240
 QY 241 GCGGGGATCTGGAGGGGGTTCGTCAAAAACCTTCCTTATCTTAAACAGCTGGCGTAACG 300
 Db |||||
 241 GCGGGGATCTGGAGGGGGTTCGTCAAAAACCTTCCTTATCTTAAACAGCTGGCGTAACG 300
 QY 301 ACAATCTGTTGTTCCTCCGGTTTTGACAAATCTGGATACACTGCGGGGACCGGATAACAG 360
 Db |||||
 301 ACAATCTGTTGTTCCTCCGGTTTTGACAAATCTGGATACACTGCGGGGACCGGATAACAG 360
 QY 361 GCCTATCACGGATCTGGACGGCGATTTTAAACAGATTCAGGAACATTTGGGAATTGG 420
 Db |||||
 361 GCCTATCACGGATCTGGACGGCGATTTTAAACAGATTCAGGAACATTTGGGAATTGG 420
 QY 421 ACCACATTTGCACAGTTGGTCAATGATGCTCACAAAACGGAATCAAGTCAATTTGTCGAC 480
 Db |||||
 421 ACCACATTTGCACAGTTGGTCAATGATGCTCACAAAACGGAATCAAGTCAATTTGTCGAC 480
 QY 481 TTTGTGCCCAATCATTCGATCTCCTTTTAAAGGCAAAACGATTCACCTTTGCGGAAGCGGC 540
 Db |||||
 481 TTTGTGCCCAATCATTCGATCTCCTTTTAAAGGCAAAACGATTCACCTTTGCGGAAGCGGC 540
 QY 541 GCCCTCTACAAATGGAACCTTATATGGGCAATTTTGTGACGCAACAAAAGGTTAC 600
 Db |||||
 541 GCCCTCTACAAATGGAACCTTATATGGGCAATTTTGTGACGCAACAAAAGGTTAC 600

QY 601 TTCCACCAATAATGGGSGCATCAGCAACTGGGACGACCGGTACGAGCGCAATGGAAAAAC 660
Db 601 TTCCACCAATAATGGGSGCATCAGCAACTGGGACGACCGGTACGAGCGCAATGGAAAAAC 660
QY 661 TTCCAGGATCCAGCCGGTTCTCGCTTGGCGGATTTGTCGAGGAAATGSCAGATTCCT 720
Db 661 TTCCAGGATCCAGCCGGTTCTCGCTTGGCGGATTTGTCGAGGAAATGSCAGATTCCT 720
QY 721 CAATACCTGACCGATGGCGGGTTCATTTGGTAGCAGATGAGCGGATTTGGCGATT 780
Db 721 CAATACCTGACCGATGGCGGGTTCATTTGGTAGCAGATGAGCGGATTTGGCGATT 780
QY 781 GATGCGGTGAGCAATTTTAATTCGGGGTTCCTCAATCGTTGGCCGATTAACCTGACCAA 840
Db 781 GATGCGGTGAGCAATTTTAATTCGGGGTTCCTCAATCGTTGGCCGATTAACCTGACCAA 840
QY 841 AAGAAGACATTTCTCGTGGGGGATGTCAGGAGATGACCCCGAAGACGCAATCAT 900
Db 841 AAGAAGACATTTCTCGTGGGGGATGTCAGGAGATGACCCCGAAGACGCAATCAT 900
QY 901 CTGGAAGAGGTCGGGTACGCCAACAACACGGGTGTCATATGTCGATTTGATCTCAAC 960
Db 901 CTGGAAGAGGTCGGGTACGCCAACAACACGGGTGTCATATGTCGATTTGATCTCAAC 960
QY 961 ACGGTATTCGAATGTTTCGGGCACATTTACGCAACGATGACGATCTTAACATATG 1020
Db 961 ACGGTATTCGAATGTTTCGGGCACATTTACGCAACGATGACGATCTTAACATATG 1020
QY 1021 GTGAACCAAAAGGGGACGAGTCAAAATACAAAGAAATCTAATCAATTTATCGATAAC 1080
Db 1021 GTGAACCAAAAGGGGACGAGTCAAAATACAAAGAAATCTAATCAATTTATCGATAAC 1080
QY 1081 CATGATATGTAAGATTTCTTCGGTAAATTCGAAACGAGGATTTTCACCGCGCTT 1140
Db 1081 CATGATATGTAAGATTTCTTCGGTAAATTCGAAACGAGGATTTTCACCGCGCTT 1140
QY 1141 GCCTTCATCTCAGTTCGGGGGTAGCCCTCCATCTATATGAAACGCAATACATG 1200
Db 1141 GCCTTCATCTCAGTTCGGGGGTAGCCCTCCATCTATATGAAACGCAATACATG 1200
QY 1201 GCAGCGGCAATGACCCGTACAAACCGGGGATGATGCCGGTTGTGATGCAACACACC 1260
Db 1201 GCAGCGGCAATGACCCGTACAAACCGGGGATGATGCCGGTTGTGATGCAACACACC 1260
QY 1261 GCCTTTAAAGAGTGTCACTCTGCGGGGTGCGCAGCAACATGCGGATCCAGTAC 1320
Db 1261 GCCTTTAAAGAGTGTCACTCTGCGGGGTGCGCAGCAACATGCGGATCCAGTAC 1320
QY 1321 GGCACCAACCCAGCGTTGGATCAACATGATGTTTACATTTATGACGAAATTTTC 1380
Db 1321 GGCACCAACCCAGCGTTGGATCAACATGATGTTTACATTTATGACGAAATTTTC 1380
QY 1381 AAGCATGTCGTTGGTGGCCATCAATCGAAACAGCAATCTCTTATTCGATTCGGT 1440
Db 1381 AAGCATGTCGTTGGTGGCCATCAATCGAAACAGCAATCTCTTATTCGATTCGGT 1440
QY 1441 TTGCAGACGCGCTTCCCAATGTCAGTATGCGGATTCCTGTCAGGCGTGTGGGGGG 1500
Db 1441 TTGCAGACGCGCTTCCCAATGTCAGTATGCGGATTCCTGTCAGGCGTGTGGGGGG 1500
QY 1501 AACGGATTTCCGTTTCCAAATGGAAGTGTGCTGTTGATCGCTTGGCGCTGAGCGGTG 1560
Db 1501 AACGGATTTCCGTTTCCAAATGGAAGTGTGCTGTTGATCGCTTGGCGCTGAGCGGTG 1560
QY 1561 TCTGTTGGCAGTACAGCAGATTCGCTTACGCGCGCAATCGGATCGTGTCTCAAT 1620
Db 1561 TCTGTTGGCAGTACAGCAGATTCGCTTACGCGCGCAATCGGATCGTGTCTCAAT 1620
QY 1621 ATGGGATTCGGGTAATGTCGTCAACATGACGAGGAAAGGTTTGGGACGACGAGGA 1680
Db 1621 ATGGGATTCGGGTAATGTCGTCAACATGACGAGGAAAGGTTTGGGACGACGAGGA 1680
QY 1681 ACCGTGACATTTGGCGGAGTGACAGCGACTGTGAAATCTCGACATCCAATCGGATTGAA 1740

RESULT 3

AAA27399 standard; DNA; 2160 BP.
XX
AC AAA27399;
XX
DT 04-AUG-2000 (first entry)
XX
DE Bacillus maltogenic alpha-amylase gene.
XX
KW Bacillus; maltogenic alpha-amylase; glucan 1,4-alpha-maltohydrolase;
KW E.C.3.2.1.133; baking; transgenic plant; Novamyl; enzyme;
KW hydrolysis; anti-staling; maltose; maltotriose; cyclodextrin; ds.
XX
OS Bacillus sp. 'NCIB 11837'.
XX
FH Key Location/Qualifiers
FT CDS 1..2160
FT /*tag= a
FT /product= "maltogenic alpha-amylase"
FT sig_peptide 1..99
FT /*tag= b
FT mat_peptide 100..2157
FT /*tag= c
FT /EC_number= "3.2.1.133"
XX
PN W0200029591-A1.
XX
PD 25-MAY-2000.
XX
PF 12-NOV-1999; 99WO-DK00624.
XX
PR 12-NOV-1998; 98DK-0001478.
PR 10-MAR-1999; 99US-0123643.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Nielsen JB, Kjaerulff S;
XX WPI; 2000-387803/33.
DR P-PSDB; AAY94271.

XX	Transgenic plant cell transformed with a nucleotide sequence encoding
PT	maltoenic alpha-amylase or beta-amylase, useful for producing
PT	transgenic wheat plant and seeds used in preparing dough and bread
PT	improver composition -
XX	Claim 5; Page 36-40; 47pp; English.
XX	The present sequence is the <i>Bacillus maltogenic</i> alpha-amylase gene. This
CC	gene is also known as glucan 1,4-alpha-maltohydrolase and Novamyl
CC	(E.C. 3.2.1.133). Maltogenic alpha-amylase hydrolyses amylose and
CC	amylpectin to maltose in the alpha-configuration, and is also able to
CC	hydrolyse maltotriose and cyclodextrin. The gene product can be used in
CC	the baking industry in flour to improve anti-staling effects i.e. to
CC	improve softness and elasticity of dough during storage. The gene can
CC	be used to produce transgenic plant seeds, especially wheat. These seeds can
CC	then be milled to produce flour which can be used to produce bread with
CC	the improved qualities.
XX	Sequence 2160 BP; 571 A; 499 C; 572 G; 518 T; 0 other;
XX	Query Match 100.0%; Score 2160; DB 21; Length 2160;
XX	Best Local Similarity 100.0%; Pred. No. 0;
XX	Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGAAAAGAAAACGCTTCTTTATTGTGGGACTGATGCTCTCATCGTCTTCTGTTTC 60
DB	1 ATGAAAAGAAAACGCTTCTTTATTGTGGGACTGATGCTCTCATCGTCTTCTGTTTC 60
QY	61 AGCGGTTCTCTTCGTACAAATCAAAACGCGCTGAGCCAGCAGTTCGCGAAGCGTCAA 120
DB	61 AGCGGTTCTCTTCGTACAAATCAAAACGCGCTGAGCCAGCAGTTCGCGAAGCGTCAA 120
QY	121 GGGGACGTGATTTACCAAGATTATCATTCACCGGTTTTACGATGGGGACACGACGAA 180
DB	121 GGGGACGTGATTTACCAAGATTATCATTCACCGGTTTTACGATGGGGACACGACGAA 180
QY	181 AATCTGCGAAAAGTTATGACATTTACGATCCGACCAATCGAAGTGAAAATGTATTGG 240
DB	181 AATCTGCGAAAAGTTATGACATTTACGATCCGACCAATCGAAGTGAAAATGTATTGG 240
QY	241 GCGGGGATCTGAGGGGGTTCGTCAAACAATCTCCTTATCTTAAACAGCTGGCGTAA 300
DB	241 GCGGGGATCTGAGGGGGTTCGTCAAACAATCTCCTTATCTTAAACAGCTGGCGTAA 300
QY	301 ACAATCTGGTTGTCCCGGTTTTGGACAAATCTGATACACTGCGGGGACCGATAACG 360
DB	301 ACAATCTGGTTGTCCCGGTTTTGGACAAATCTGATACACTGCGGGGACCGATAACG 360
QY	361 GGCTATCAGGATACCTGACGGCGCATTTTAAACAGATTTGAGGACATTTTCGGGA 420
DB	361 GGCTATCAGGATACCTGACGGCGCATTTTAAACAGATTTGAGGACATTTTCGGGA 420
QY	421 ACCACATTTGACACGTTGGTCAATGATCTCACCAAAACGGAATCAAGGTGATTTGCG 480
DB	421 ACCACATTTGACACGTTGGTCAATGATCTCACCAAAACGGAATCAAGGTGATTTGCG 480
QY	481 TTTGTGCCAATCATTCGACTCCTTTTAAGCAAAACGANTTCCACTTTTCGGGAAGCG 540
DB	481 TTTGTGCCAATCATTCGACTCCTTTTAAGCAAAACGANTTCCACTTTTCGGGAAGCG 540
QY	541 GCCTCTTACACAAATGAAACCTTATTTGGGGAATTTATTTTATGACGCAACAAAGG 600
DB	541 GCCTCTTACACAAATGAAACCTTATTTGGGGAATTTATTTTATGACGCAACAAAGG 600
QY	601 TTCCACCAATAATGGGACATCAAGCAACTGGGACACCGGTACGAGCGCAATGGAAA 660
DB	601 TTCCACCAATAATGGGACATCAAGCAACTGGGACACCGGTACGAGCGCAATGGAAA 660
QY	661 TTACCGGATACAGCGGTTTCTCGCTTCGCAATTTGTTCGCAAGAAATGGCACGAT 720
DB	661 TTACCGGATACAGCGGTTTCTCGCTTCGCAATTTGTTCGCAAGAAATGGCACGAT 720

Db 1801 TCCAGCAATCTGTAATCTTACATATTTTGAAGTGAACGACATCGTGTGTTTACT 1860
Qy 1861 GTGAAAGTGCCTCCGACCACTGGGGGATAAGATTACCTGAGCGCAACATACG 1920
Db 1861 GTGAAAGTGCCTCCGACCACTGGGGGATAAGATTACCTGAGCGCAACATACG 1920
Qy 1921 GAATTGGGAATTGGAGCAGGATACGAGCGGCGGTTTAACTTGGCGAGGCGCCCTG 1980
Db 1921 GAATTGGGAATTGGAGCAGGATACGAGCGGCGGTTTAACTTGGCGAGGCGCCCTG 1980
Qy 1981 CTCGCGCCCAATATCCGAGTTGTTTATGTTATTCAGGTTCCAGCAGGAAGAGATT 2040
Db 1981 CTCGCGCCCAATATCCGAGTTGTTTATGTTATTCAGGTTCCAGCAGGAAGAGATT 2040
Qy 2041 CAATTCAAGTTCTTCATCAACGCTGCGATGGAACGANTCAATGGGAGATGTTTCAAC 2100
Db 2041 CAATTCAAGTTCTTCATCAACGCTGCGATGGAACGANTCAATGGGAGATGTTTCAAC 2100
Qy 2101 CAGTGGCCCAACTCCCAACGCGGTGCAACCGGTTAAGTTTACGTTGGCGAAAACTAG 2160
Db 2101 CAGTGGCCCAACTCCCAACGCGGTGCAACCGGTTAAGTTTACGTTGGCGAAAACTAG 2160

RESULT 4

AAPE2389

ID AAPE2389 standard; DNA; 2160 BP.

XX

AC AAPE2389;

XX

DT 05-NOV-2001 (first entry)

XX

DE Bacillus maltogenic amylase coding sequence.

XX

KW Maltogenic amylase; mutant; modified starch; food; cosmetics; paint;
KW pharmaceutical; polymer; textile; leather processing; ore flotation;
KW oil drilling; agricultural spraying; ds.

XX

OS Bacillus sp.

XX

FH Key Location/Qualifiers

FT CDS

1..2160

/*tag= a

/product= "maltogenic amylase"

XX

PN WO200116348-A1.

XX

PD 08-MAR-2001.

XX

XX 21-AUG-2000; 2000WO-DK00460.

XX

XX 01-SEP-1999; 99DK-0001219.

XX

XX 01-SEP-1999; 99US-0151847.

XX

XX (NOVO) NOVOZYMES AS.

XX

PI Pedersen S, Vang Hendriksen H;

XX

XX WPI; 2001-273341/28.

XX

XX P-PSDB; AAG59956.

XX

PT Preparing enzymatically modified starch useful for e.g. incorporation

XX

PT in emulsions and food products involves treating starch derivative with

XX

PT maltogenic amylase which retains amylase activity in low pH range

XX

XX Claim 1; Page 61-65; 77pp; English.

XX

XX The present invention describes a method of preparing enzymatically

XX

XX modified starch derivatives, involving treating the starch derivative

XX

XX with a maltogenic amylase and recovering the modified starch derivative.

XX

XX They are useful as emulsifying agents in food, in the cosmetics, paint,

XX

XX pharmaceutical and polymer industries, in textiles, leather processing,

XX

XX ore flotation, oil drilling and agricultural spraying. The present

CC sequence is the coding sequence of a maltogenic amylase from Bacillus.

XX

SQ Sequence 2160 BP; 571 A; 499 C; 572 G; 518 T; 0 other;

XX

Query Match 100.0%; Score 2160; DB 22; Length 2160;

XX

Best Local Similarity 100.0%; Pred. No. 0;

XX

Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

Qy 1 ATGAAAAAGAAAAACGCTTCTTATTTCTGGGACTGATGCTCTCATCGGTTCTTCTTC 60

Db

1 ATGAAAAAGAAAAACGCTTCTTATTTCTGGGACTGATGCTCTCATCGGTTCTTCTTC 60

Qy

61 AGCGTTCTCTTCGTTACAAATCCAAACGCGCTGAAGCCAGAGTTCCGCAAGCGTCAAA 120

Db

61 AGCGTTCTCTTCGTTACAAATCCAAACGCGCTGAAGCCAGAGTTCCGCAAGCGTCAAA 120

Qy

121 GGGGACGTGATTTACAGATTTATCTTACCGGTTTACGATGGGACACGACGACCAAC 180

Db

121 GGGGACGTGATTTACAGATTTATCTTACCGGTTTACGATGGGACACGACGACCAAC 180

Qy

181 AATCTGCAAAAAGTTATGAGCTTTACGATCCGACCAATCGAAGTGGAAAAATGTATTGG 240

Db

181 AATCTGCAAAAAGTTATGAGCTTTACGATCCGACCAATCGAAGTGGAAAAATGTATTGG 240

Qy

241 GGGGGGATCTGGAGGGGTTTGGACAACTCTGGATACACTGGCGGACCCGATAACACG 360

Db

241 GGGGGGATCTGGAGGGGTTTGGACAACTCTGGATACACTGGCGGACCCGATAACACG 360

Qy

301 ACAATCTGTTTGTCCCGGTTTGGACAACTCTGGATACACTGGCGGACCCGATAACACG 360

Db

301 ACAATCTGTTTGTCCCGGTTTGGACAACTCTGGATACACTGGCGGACCCGATAACACG 360

Qy

361 GCGTATCGAGTACTGGACGCGGATTTTAAACAGATTGAGGAAATGTTGGAAATTGG 420

Db

361 GCGTATCGAGTACTGGACGCGGATTTTAAACAGATTGAGGAAATGTTGGAAATTGG 420

Qy

421 ACCAATTTGACAGTTGGTCAATGATGCTCACCACGAGGATCAAGTGTATGTCGAC 480

Db

421 ACCAATTTGACAGTTGGTCAATGATGCTCACCACGAGGATCAAGTGTATGTCGAC 480

Qy

481 TTGTGCCCCAATCATTCGACTCCTTTTAAAGCAACAGATTCCACCTTTCCGGAAGCGGC 540

Db

481 TTGTGCCCCAATCATTCGACTCCTTTTAAAGCAACAGATTCCACCTTTCCGGAAGCGGC 540

Qy

541 GCGCTCTACACATGAACTATATGGGCAATTTTATGATGACGCAACAAAGGTTAC 600

Db

541 GCGCTCTACACATGAACTATATGGGCAATTTTATGATGACGCAACAAAGGTTAC 600

Qy

601 TTCCACCAATATGGGGACATCACTGGAACCTGGGACGCGGTACGAGCGCAATGGAAAAAC 660

Db

601 TTCCACCAATATGGGGACATCACTGGAACCTGGGACGCGGTACGAGCGCAATGGAAAAAC 660

Qy

661 TTACGAGATCCAGCGGTTTCTGCTTCCGATTTGTCGAGGAAAAATGGACAGATTGCT 720

Db

661 TTACGAGATCCAGCGGTTTCTGCTTCCGATTTGTCGAGGAAAAATGGACAGATTGCT 720

Qy

721 CAATACCTGACCGGTTTCTGCTTCCGATTTGTCGAGGAAAAATGGACAGATTGCT 780

Db

721 CAATACCTGACCGGTTTCTGCTTCCGATTTGTCGAGGAAAAATGGACAGATTGCT 780

Qy

781 GATCGGTGAAGCAATTTTAAATCGGGTTCTCCAAATCGTTGGCCGATAAACTGTACCAA 840

Db

781 GATCGGTGAAGCAATTTTAAATCGGGTTCTCCAAATCGTTGGCCGATAAACTGTACCAA 840

Qy

841 AAGAAGACATTTTCTGTTGGGGGATGTTACGAGATGACCCCGGAACAGCAATCAT 900

Db

841 AAGAAGACATTTTCTGTTGGGGGATGTTACGAGATGACCCCGGAACAGCAATCAT 900

Qy

901 CTGGAAGAGTCCCGTACGCAACACGCGTCTCAATGCTGCAATTTTCAATCTCAAC 960

Db

901 CTGGAAGAGTCCCGTACGCAACACGCGTCTCAATGCTGCAATTTTCAATCTCAAC 960

Qy

961 ACGGTGATTGCAAAATGTTTGGCACTTTACGCAACGATGTACGATCTTAACAATATG 1020

Db 961 ACAGTATGTCAGATGTTGTCGGCACAATTTACGCAACGATGTCAGATTTTACATATG 1020
QY 1021 GTGAACAAACGGGGAAGAGTACAATAACAAGAAAACTTAATCAATTTATCGATAAC 1080
Db 1021 GTGAACAAACGGGGAAGAGTACAATAACAAGAAAACTTAATCAATTTATCGATAAC 1080
QY 1081 CATGATATGTCAGATTTCTTTCGGTAAATTCGAACAAGGGGAATTTGACACGGGGCTT 1140
Db 1081 CATGATATGTCAGATTTCTTTCGGTAAATTCGAACAAGGGGAATTTGACACGGGGCTT 1140
QY 1141 GCTTTTCATTTCTCACTTCGGGGGTACGCCCTCCATCTATTATGAAACCAATACATG 1200
Db 1141 GCTTTTCATTTCTCACTTCGGGGGTACGCCCTCCATCTATTATGAAACCAATACATG 1200
QY 1201 GCAGGCGCAATGACCCGTCACAACGGGGGATGATGCGGGTTTGATACGACCAACACC 1260
Db 1201 GCAGGCGCAATGACCCGTCACAACGGGGGATGATGCGGGTTTGATACGACCAACACC 1260
QY 1261 GCCTTTAAAGAGTGTCACTCTGCGGGGTTCGCGAGGAACATGCGGATCCAGTAC 1320
Db 1261 GCCTTTAAAGAGTGTCACTCTGCGGGGTTCGCGAGGAACATGCGGATCCAGTAC 1320
QY 1321 GGCACCAACCCAGCGTGTGATCAACAATGATGTTTACATTTATGAACGGAATTTTTC 1380
Db 1321 GGCACCAACCCAGCGTGTGATCAACAATGATGTTTACATTTATGAACGGAATTTTTC 1380
QY 1381 AAGCATGTGCTTGTGGTGCATCAATCGAAACAGCAATCTCTTTCGATTCGGGT 1440
Db 1381 AAGCATGTGCTTGTGGTGCATCAATCGAAACAGCAATCTCTTTCGATTCGGGT 1440
QY 1441 TTGCACAGCGCTTCGCAATGCGAGTATGCGGATATCTCTCAGGGCTGTGGGGGG 1500
Db 1441 TTGCACAGCGCTTCGCAATGCGAGTATGCGGATATCTCTCAGGGCTGTGGGGGG 1500
QY 1501 AACGGGATTCGGTTTCCAAATGGAAGTGTGCTTCGCTTCAGCTTGCGCTGAGCCGTG 1560
Db 1501 AACGGGATTCGGTTTCCAAATGGAAGTGTGCTTCGCTTCAGCTTGCGCTGAGCCGTG 1560
QY 1561 TCTGTTGGCAGTACAGACATCCGCTTCAGCGCGCAATCGGATCGGTGTCCTCAAT 1620
Db 1561 TCTGTTGGCAGTACAGACATCCGCTTCAGCGCGCAATCGGATCGGTGTCCTCAAT 1620
QY 1621 ATGGGATTCGGGTAATGCTGTCAGATCGAGCGGGAAGGTTTGGAGCGACGAGGA 1680
Db 1621 ATGGGATTCGGGTAATGCTGTCAGATCGAGCGGGAAGGTTTGGAGCGACGAGGA 1680
QY 1681 ACCGTGACATTTGGCGGAGTGACAGCGACTGTGAATCTCTGACATCCAAATCGGATGAA 1740
Db 1681 ACCGTGACATTTGGCGGAGTGACAGCGACTGTGAATCTCTGACATCCAAATCGGATGAA 1740
QY 1741 GTGTAGCTTCCCAATGCGCGCGGCTGACCGATGTGAAGTCAACCGGGGTGAGTT 1800
Db 1741 GTGTAGCTTCCCAATGCGCGCGGCTGACCGATGTGAAGTCAACCGGGGTGAGTT 1800
QY 1801 TCAGCAATCTCTATTCTTACAAATTTTGAAGTGAACGAGACATCGGTGTGTTACT 1860
Db 1801 TCAGCAATCTCTATTCTTACAAATTTTGAAGTGAACGAGACATCGGTGTGTTACT 1860
QY 1861 GTGAAGGTGCGCTTCGACCAACCTTGGGGATPAAGATTTACCTGACGGGCAATACCG 1920
Db 1861 GTGAAGGTGCGCTTCGACCAACCTTGGGGATPAAGATTTACCTGACGGGCAATACCG 1920
QY 1921 GAATTTGGGAATTTGAGCAGCAGATACGAGCGGCGCTTAATGATGCGAGGGCCCTG 1980
Db 1921 GAATTTGGGAATTTGAGCAGCAGATACGAGCGGCGCTTAATGATGCGAGGGCCCTG 1980
QY 1981 CTGCGGCCAATTTATCCGGATTTGTTTATGATTCAGGTTTCAGAGGAAAGAGAT 2040
Db 1981 CTGCGGCCAATTTATCCGGATTTGTTTATGATTCAGGTTTCAGAGGAAAGAGAT 2040
QY 2041 CAATTCAGTCTTCATCAACGCTGCGATGGAACGATTCATATGGGAGATGTTTCGAAC 2100

Db 2041 CAATTCAGTCTTCATCAACGCTGCGATGGAACGATTCATGGGAGATGTTTCGAAC 2100
QY 2101 CACGTGGCCCAACTCCACGGGTGCAACCGGTAACTTACTGTTACGTGGCAAACTAG 2160
Db 2101 CACGTGGCCCAACTCCACGGGTGCAACCGGTAACTTACTGTTACGTGGCAAACTAG 2160
RESULT 5
AAF77165
ID AAF77165 standard; DNA; 2160 BP.
XX
AC AAF77165;
XX
DT 23-MAY-2001 (first entry)
XX
DE Bacillus maltogenic amylase DNA.
XX
KW Maltose; starch; syrup; fat substitute; mutant; mutein; ds.
OS
XX Bacillus sp.
XX
PN WO200116349-A1.
XX
PD 08-MAR-2001.
XX
PF 21-AUG-2000; 2000WO-DK00461.
XX
PR 01-SEP-1999; 99DK-0001220.
XX
PR 12-JAN-2000; 2000DK-0000035.
XX
PA (NOVO) NOVOZYMES AS.
XX
PI Pedersen S, Vang Hendriksen H;
XX
DR WPI; 2001-257704/26.
XX
PT Preparation of maltose and modified starch, useful e.g. for paper
coating and in food processing, by treating starch with modified
PT Bacillus maltogenic amylase
XX
PS Claim 1; Page 82-86; 99pp; English.
XX
CC The present invention relates to preparation of maltose and/or
modified starch by treating starch with a variant of a maltogenic
CC amylase. The method is used to produce high or low maltose
CC syrups or speciality syrups, useful e.g. in baking and brewing. Also
CC used to make starch for use in coating/sizing paper and in food products
CC (beverages, beverage flavour concentrates and flavouring agents), as a
CC fat substitute and to make maltose for use e.g. in intravenous feeding
CC solutions or as intermediate for the sweetener maltitol.
XX
SQ Sequence 2160 BP; 571 A; 499 C; 572 G; 518 T; 0 other;
Query Match 100.0%; Score 2160; DB 22; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAGAAAACGCTTTCTTTATTTGAGGACTGATGCTCTCATCGGTCTTCTGTC 60
Db 1 ATGAAAAGAAAACGCTTTCTTTATTTGAGGACTGATGCTCTCATCGGTCTTCTGTC 60
QY 61 AGCGGTTCTCTTCGTACAAATCCAAACCGCGTGAAGCCAGAGTTCCGCAAGCGTCAAA 120
Db 61 AGCGGTTCTCTTCGTACAAATCCAAACCGCGTGAAGCCAGAGTTCCGCAAGCGTCAAA 120
QY 121 GGGGCGGTGATTTACCAAGATTTATGATGACCGGTTTTACGATGGGACACGACGAACAC 180
Db 121 GGGGCGGTGATTTACCAAGATTTATGATGACCGGTTTTACGATGGGACACGACGAACAC 180
QY 181 ATCTCTGCCAAAAGTTATGACTTTACGATCCGACCAATCGAAGTGGAAATGTATTGG 240
Db 181 ATCTCTGCCAAAAGTTATGACTTTACGATCCGACCAATCGAAGTGGAAATGTATTGG 240

QY 241 GCGGGGATCTGGAGGGGGTTCGTCAAAAACTTCCTTAATCTTAACAGCTGGGGTAAAC 300
DB 241 GCGGGGATCTGGAGGGGGTTCGTCAAAAACTTCCTTAATCTTAACAGCTGGGGTAAAC 300
QY 301 ACAATCTGGTTCCTCCCGGTTTGACAACTCTGATACACTGGCGGACCGATAACACG 360
DB 301 ACAATCTGGTTCCTCCCGGTTTGACAACTCTGATACACTGGCGGACCGATAACACG 360
QY 361 GCCTATCAAGGATCTGACGCGCGGATTTAAACAGAGATGAGGAACATTTGGGAATGG 420
DB 361 GCCTATCAAGGATCTGACGCGCGGATTTAAACAGAGATGAGGAACATTTGGGAATGG 420
QY 421 ACCAATTTGACGCTTGGTCAATGATGCTCAACAAAACGGAATCAAGGATTTGTCGAC 480
DB 421 ACCAATTTGACGCTTGGTCAATGATGCTCAACAAAACGGAATCAAGGATTTGTCGAC 480
QY 481 TTTGTGCCAATCACTCGACTCTTTTAAAGGCAACGATTCACCTTTGCGGAGGCGGC 540
DB 481 TTTGTGCCAATCACTCGACTCTTTTAAAGGCAACGATTCACCTTTGCGGAGGCGGC 540
QY 541 GGCCTCTACAAATGGAACCTATATGGGCAATTTATTTGATGACGCAACAAAAGGCTAC 600
DB 541 GGCCTCTACAAATGGAACCTATATGGGCAATTTATTTGATGACGCAACAAAAGGCTAC 600
QY 601 TTCCACCAATATGGGACATCAGCAACTGGGACACCGGTACGAGCGCGCAATGGAAAC 660
DB 601 TTCCACCAATATGGGACATCAGCAACTGGGACACCGGTACGAGCGCGCAATGGAAAC 660
QY 661 TTCCAGGATCCAGCGGTTTCTCGGTTGCGGATTTGTCGAGGAAATGGACGATTCCT 720
DB 661 TTCCAGGATCCAGCGGTTTCTCGGTTGCGGATTTGTCGAGGAAATGGACGATTCCT 720
QY 721 CAATACCTGACCGATGCGGCTTCAATTTGGTAGCAATGAGCGCGATGTTGCGGAT 780
DB 721 CAATACCTGACCGATGCGGCTTCAATTTGGTAGCAATGAGCGCGATGTTGCGGAT 780
QY 781 GATGCGGTGAAGCAATTTAATTCGGGTTTCTCAAAATGCTTGCAGTAACTGTACCA 840
DB 781 GATGCGGTGAAGCAATTTAATTCGGGTTTCTCAAAATGCTTGCAGTAACTGTACCA 840
QY 841 AAGAAAGACATTTTCTGTTGGGGAATGTTACGAGATGACCCCGGAACAGCCAAATCAT 900
DB 841 AAGAAAGACATTTTCTGTTGGGGAATGTTACGAGATGACCCCGGAACAGCCAAATCAT 900
QY 901 CTGGAAGAGTCCGATGACCAACACAGCGGTGCTAATGCTGATTTGATCTCAAC 960
DB 901 CTGGAAGAGTCCGATGACCAACACAGCGGTGCTAATGCTGATTTGATCTCAAC 960
QY 961 ACGGTGATTCGAATGTTGTCGGCAATTTACGCAACGATGTACGATCTTAACAAATATG 1020
DB 961 ACGGTGATTCGAATGTTGTCGGCAATTTACGCAACGATGTACGATCTTAACAAATATG 1020
QY 1021 GTGAACCAACGCGGAACGATGTAACAAATCAAGAAATCTAATCAATTTATGATTAAC 1080
DB 1021 GTGAACCAACGCGGAACGATGTAACAAATCAAGAAATCTAATCAATTTATGATTAAC 1080
QY 1081 CATGATATGCAAGATTTCTTCGGTAAATTCGCAACGAGGATTTGACCAACGCGCTT 1140
DB 1081 CATGATATGCAAGATTTCTTCGGTAAATTCGCAACGAGGATTTGACCAACGCGCTT 1140
QY 1141 GCTTTCAATCTCACTTCGCGGGTACGCGCTTCAATCTTATGGAACGAAATATGATG 1200
DB 1141 GCTTTCAATCTCACTTCGCGGGTACGCGCTTCAATCTTATGGAACGAAATATGATG 1200
QY 1201 GAGGCGGATGACCGGTACACCGGGGATGATGCGGCTTTGATACGAACACACC 1260
DB 1201 GAGGCGGATGACCGGTACACCGGGGATGATGCGGCTTTGATACGAACACACC 1260
QY 1261 GCTTTAAGAGGTGCACTCTGCGGGTTGCGCAAGCAATGCGCGATCCAGTAC 1320
DB 1261 GCTTTAAGAGGTGCACTCTGCGGGTTGCGCAAGCAATGCGCGATCCAGTAC 1320
QY 1321 GGCACCACCCAGCGTTGGATCAACAAATGATGTTTATGATTAACGGAATTTTTC 1380

DB 1321 GGCACCACCCAGCGTTGGATCAACAAATGATGTTTATGATTAACGGAATTTTTC 1380
QY 1381 AAGCATGCTGCTGCTGGCCATCAATCGAACAACGCAATCTCTTATTCGATTTCCGGT 1440
DB 1381 AAGCATGCTGCTGCTGGCCATCAATCGAACAACGCAATCTCTTATTCGATTTCCGGT 1440
QY 1441 TTGCAAGCGCTTCCCAATGCGAGCTATGCGGATTAATCTGTCAGGGCTGTTGGGGGG 1500
DB 1441 TTGCAAGCGCTTCCCAATGCGAGCTATGCGGATTAATCTGTCAGGGCTGTTGGGGGG 1500
QY 1501 AACGGATTTCCGTTTCAATGGAAGTCTGCTTCTGTCACGCTTGGCTCGAGCCGTTG 1560
DB 1501 AACGGATTTCCGTTTCAATGGAAGTCTGCTTCTGTCACGCTTGGCTCGAGCCGTTG 1560
QY 1561 TCTGTTGGCAGTACAGCACATCCGTTTCAGCGCGCAAAATCGGATCGTTGCTCCAAAT 1620
DB 1561 TCTGTTGGCAGTACAGCACATCCGTTTCAGCGCGCAAAATCGGATCGTTGCTCCAAAT 1620
QY 1621 ATGGGATTTCCGGTAAATGCTCAOGATCGAGGAAGGTTTGGGACGACGCAAGGA 1680
DB 1621 ATGGGATTTCCGGTAAATGCTCAOGATCGAGGAAGGTTTGGGACGACGCAAGGA 1680
QY 1681 ACCGTGACATTTGGCGGAGTACAGCACTGTAATCTCGAATCCGATCCGATCGATTGAA 1740
DB 1681 ACCGTGACATTTGGCGGAGTACAGCACTGTAATCTCGAATCCGATCCGATCGATTGAA 1740
QY 1741 GTGTACGTTCCCAACATGCGCGCGCTGACCGATGTGAAGTCAACGCGGCTGGAGTT 1800
DB 1741 GTGTACGTTCCCAACATGCGCGCGCTGACCGATGTGAAGTCAACGCGGCTGGAGTT 1800
QY 1801 TCCAGCAATCTGATTTCTTACAAATTTTGAAGTGAAGCAATCGGTTGTTTACT 1860
DB 1801 TCCAGCAATCTGATTTCTTACAAATTTTGAAGTGAAGCAATCGGTTGTTTACT 1860
QY 1861 GTGAAAAGTCCGCTCCGACCACTCGGGGATAAGATTACTGACGGGCAACATACCG 1920
DB 1861 GTGAAAAGTCCGCTCCGACCACTCGGGGATAAGATTACTGACGGGCAACATACCG 1920
QY 1921 GAATGGGAATTTGAGCACGATACGAGCGGACCGTTAAATGCGAGGCGCCCTG 1980
DB 1921 GAATGGGAATTTGAGCACGATACGAGCGGACCGTTAAATGCGAGGCGCCCTG 1980
QY 1981 CTCGCGCCCAATTTCCGATTTGTTTATGATTCAGCGTTCCAGCAAGAAAGATG 2040
DB 1981 CTCGCGCCCAATTTCCGATTTGTTTATGATTCAGCGTTCCAGCAAGAAAGATG 2040
QY 2041 CAATTCAGTCTTCAATCAAGCGTTCGAGTGAACGATTAATGGAAGATGTTGGAAC 2100
DB 2041 CAATTCAGTCTTCAATCAAGCGTTCGAGTGAACGATTAATGGAAGATGTTGGAAC 2100
QY 2101 CAGTGGCAACTCTCCACGCGGTGCAACCGGTAACTTACTGTTAGTGGCAAACTAG 2160
DB 2101 CAGTGGCAACTCTCCACGCGGTGCAACCGGTAACTTACTGTTAGTGGCAAACTAG 2160

RESULT 6

AAV28880
ID AAV28880 standard; DNA; 576 BP.

XX AAV28880;

DT 09-SEP-1998 (first entry)

XX Nucleotide sequence of the starch binding domain DB-domain.
XX Starch binding domain; SBD; DB-domain; oral; dental plaque;
XX toothpaste; dental cream, mouthwash; ss.

OS Bacillus steatothermophilus.

XX Key
FH Location/Qualifiers
FT CDS 1..576


```

1822 AATATTGAGTGAACGACGATCGGTGTGTTTACTGTGAAAAGTGGCCTCCGACC 1881
1823 AATATTGAGTGAACGACGATCGGTGTGTTTACTGTGAAAAGTGGCCTCCGACC 300
1882 AACCTGGGGGATAGATTACCTGACGGGACACATACCGAATTTGGGAAATGGAGCAG 1941
301 AACCTGGGGGATAGATTACCTGACGGGACACATACCGAATTTGGGAAATGGAGCAG 360
1942 GATACGACGAGCGGTATCAATGCGCAAGGCGCCCTGCTCGCGCCCAATTTCCGGAT 2001
361 GATACGACGAGCGGTATCAATGCGCAAGGCGCCCTGCTCGCGCCCAATTTCCGGAT 420
2002 TGGTTTATGATTCACGGTTCACGAGGAAGACGATCAATTCAGTTCTTATCAAG 2061
421 TGGTTTATGATTCACGGTTCACGAGGAAGACGATCAATTCAGTTCTTATCAAG 480
2062 CTTGGGATGGAAGCAATCAATGGGAGATGTTGCAACCACTGGCCCAACTCCCAAG 2121
481 CTTGGGATGGAAGCAATCAATGGGAGATGTTGCAACCACTGGCCCAACTCCCAAG 540
2122 GGTGCAACCGGTAAACATTAATGTTAGTGGCAAAAC 2157
541 GGTGCAACCGGTAAACATTAATGTTAGTGGCAAAAC 576

RESULT 8
ID AAN60739 standard; DNA; 2058 BP.
XX AC AAN60739;
XX DT 25-MAR-2003 (updated)
XX DT 08-AUG-1991 (first entry)
XX DE Sequence encoding cyclomaltodextrin glucanotransferase (CGTase).
XX KW Enzyme; cyclodextrin; glycosyl-sucrose; ss.
XX OS Bacillus macerans.
XX FH Key
XX FT Location/Qualifiers
XX FT 1..2058
XX FT /*tag= a
XX FT
XX FN FR2574081-A.
XX PD 06-JUN-1986.
XX PF 03-DEC-1985; 85FR-0017835.
XX PR 14-OCT-1985; 85JP-0228169.
XX PR 03-DEC-1984; 84JP-0255339.
XX PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX PI Sugimoto T, Kubota M, Sakai S;
XX WP; 1986-184710/29.
XX DR P-PSDB; AAP61518.
XX PT New polypeptide(s) - with cyclomaltodextrin glucan:transferase
XX PT activity, produced by recombinant DNA techniques
XX PS Example; Table 4-1, Page 33-36; 54pp; French.
XX CC A polypeptide comprising 680 AA units, corresp. to CGTase produced
XX CC by Bacillus stearotherophilus, and a polypeptide comprising 686 AA
XX CC units, corresp. to CGTase produced by B.macerans are claimed. Also
XX CC claimed are polypeptides which comprise one or more of AAP60813-P60817,
XX CC with a signal sequence (AAP60818), N-terminal sequence (AAP60819) and
XX CC transit peptide (AAP60820).
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX CC (Updated on 25-MAR-2003 to correct PI field.)

```

```

XX SQ Sequence 2058 BP; 517 A; 635 C; 523 G; 383 T; 0 other;
Query Match 16.1%; Score 348.8; DB 7; Length 2058;
Best Local Similarity 52.3%; Pred. No. 1.4e-99;
Matches 1078; Conservative 0; Mismatches 912; Indels 72; Gaps 11;
QY 123 GGACGTGATTTACCAAGATTATCATTTGACCGGTTTACGATGGGACACGACAAACAA 182
DB 42 GGAATAGGTTTACCAAGATTATTAACCGACCGGTTTGTGGACG---CAATTCGCCCAACA 98
QY 183 TCCTGCCAAAGATTATGACGTTTACGATCCGACCAATCAAGTGGAAATGTTATTTGGGG 242
DB 99 CCCGACCGGAGAGCGCTTCAGACGCGATCATTCACACCTGAAGCTGTTATTCGGGGGCGA 158
QY 243 CGGGGATCTGGAGGGGGTTCGTCAAAAACCTTCCTATCTTAAACAGCTGGGCGTAAACGAC 302
DB 159 CTGCGAGGGGATCAGCAAAATCAACGACGCGCTATCTGACCGGAATGGGCATCACCGC 218
QY 303 AATCTGTTGTCCTCCGGTTTGGACAATCTGGAATACACT-----GGCGGGCAACCGA 353
DB 219 CCTCTGGATCTCGACCGCGTTGAGAAACATCACCGCGCTCATCAATATTATTCGGCGCTCAA 278
QY 354 TAACACGGGCTATCAGGATACGACGCGGATTTTAAACGATTTGAGGAACATTTCCG 413
DB 279 CAATACAGCTTACCAAGGTTACTGGCTCTCGGACTTCAAGAGACCAATGCCGCTTCGG 338
QY 414 GAATTGACCAACATTTGACGCTTGGTCAATGATGCTACCAAAACGGAATCAAGGTGAT 473
DB 339 CAGCTTCAACCGCTTCTCAATTTGATCGCGCGAGCGAATTCACACAATATCAAGGTAGT 398
QY 474 TGTGACTTTGGGCGCAATCATTCGACTCTCTTTAAGGCAACGATTCACCTTTTGGGGA 533
DB 399 TATGACTTTTGCACCTTAATCAACCAACCGCGTTTCGATACGGHCCCTCTGTTGCCCGA 458
QY 534 AGGCGGCGCTCTCAACAATGGAACCTATATGGCAATTTATTTGATGACGCAACAA 593
DB 459 GAACGGCGCGCTCTACACAACGGAACGCTGCTCGGCAAGTA---TAGCAACGATACCGC 515
QY 594 AGGTAATCTCCACCAATATGAGGACATCAGCAACTGGGACACCGGTACGAGCGCAATG 653
DB 516 CGGCTGTTCCACCAATGCGCGCACCGATTTCT-----CGACGAC 557
QY 654 GAATAACTTCACGATCCAGCGGTTTCTGCTTGGGATTTGTCGAGGAATGGCAC 713
DB 558 TGAAGCGGTATCTCAAGAACCTGTACGATCTCGGGATATCAATCAGAACACAAACAC 617
QY 714 GATTGCTCAATACCTGACCGATGGCGGTTCAATTTGGTAGCATGAGCGGATGTTT 773
DB 618 CATCGACTCGTATCTCAAGGAATCGATCCAGCTGTGCTGATCTCGGAGTCGACGGCAT 677
QY 774 GGGATTTGATGGGTGAGCAATTTAATTCGGGTTTCTCCAAATCGTTGGCCGATAACT 833
DB 678 CCGCTTCGACCGCGTGAAGCATATGCTTCAGGCTCGGACAGAGTACGTTCTGTCGAT 737
QY 834 GTACCAAAA---AGAAAGACATTTTCTGTTGGGGAATGTTACGGAGATGACCCCGAAC 890
DB 738 CTACGACGCGCCCATCCCGGTGTTTCACTTCGGTGAATGTTTCTCGGCCCGCGACGAAT 797
QY 891 AGCCAATCAPCTGGAAAAGGTCGGGTACGCCCAACACAGCGGTGTCAATGTGTGGATTT 950
DB 798 GAAC-----CAGGACAAACATCAACTTCGGGAATCAGACGCGCATGCACCTGTGGACTT 851
QY 951 TGATCTCAACACCGGTATTCGAATGTTTGGCAATTTAGCGAAACGATGTACGATCT 1010
DB 852 TGGCTTTGCGCAGGAATCCGTTGAAGTTTTCGCGCAAAAGTCGGGACGATGACCGACCT 911
1011 TAACAATATGTTGAACCAACGAGGGAACGATGTAATACAAAGAAATCTAATCACTT 1070
DB 912 GAATCTCGTGAATCTCAGCACCGGCTCAGCTATAATTCATCAACACATGTTAGCTT 971
1071 TATCGATAACCATGATGATGTCAGATTTCTTTCGGTAAATTCGAACAGGCAATTTGCA 1130

```



```
Db 972 CATCGACACCATGATCATGGACCGCTTCCAGACAGCGGAGCGAGCACTCGCCCGACCGA 1031
QY 1131 CAGAGCGCTTGTCTTTCATCTCACTTCGGGGGTAGCGCTCCATCTATTATGGAACCGA 1190
Db 1032 GAGGCTCTTGGGTAAGCGCTGACTTCCGGCGGCTTCGGGCACTCTACTACGGTACAGA 1091
QY 1191 ACATATCATGGAGCGGCAATGACCGGTACAAACCGGGGATGATCGCGGCTTTGATAC 1250
Db 1092 GCAATATATGACCGGCAACGGCGACCGCAACACCGCGCATGATGACCGGCTTCGATAC 1151
QY 1251 GACACACCGCTTTTAAAGAGTGTCACTCTGGGGSETTGGCAGGAACAATGCCGCG 1310
Db 1152 GACACAGACAGGTCACAAAGTGATCAACGGCTGCTCGCTTCGCAAGTCCAACCCGCG 1211
QY 1311 GATCCAGTACGGCACCAACCCAGCGTTGGATCAACAATGATGTTTATATTTATGAACG 1370
Db 1212 TCTCGCTACGCTCGACAGCCAGCGTTGGGTGAACAGCGAGCTTACGTATATGAACG 1271
QY 1371 GAATTTTCAACGATGTCGTGTGGTGGCCATCAATGAAACAGCAATCTCCCTATTC 1430
Db 1272 CAAGTTTCGGAGCAACGTAGCTCTCGTTGCCGTCAACCGCAGCTCGACGACTGCCCTATCC 1331
QY 1431 GATTTCGGGTTTCAGAGCGCTTGCACAAATGCGACTATGCGAATTATCTGTCAGGGCT 1490
Db 1332 GATATCGGAGCGCTTACTGTCTGCGCAACGGAAGTATACCGAGCTTCTCGCGGCT 1391
QY 1491 GTTGGGGGGAACGGGATTTCCGTTTCCAAATGAAGTGTGCTGTTCAAGTTTCGCGC 1550
Db 1392 GCTTAATGGCAATTCATTAACGTTAAACGGCGGCAAGTCAACAACTTTACACTTGCACG 1451
QY 1551 TGGAGCGGTGTCTGTTTGGCAGTACAGCATCCGCTTCAGCGCGCAATCCGATCGT 1610
Db 1452 GGGCGGTACGGGAGTCTGGCAGTACAGCAGCGGAATCTCGCGAATTATCGCGAAGT 1511
QY 1611 TGCTCCAAATATGGGGAATTCGGGTAAATGTGTGTCAGATCGACGGGAAAGTTTGGGAC 1670
Db 1512 CGGCCGACTATGGGCAGCGCGGCAACACCATCAATGATCGAGCGAGCGGCTTCGGTAC 1571
QY 1671 GACCGAGGGAACCGTGACATTTGG-----CGAGTGAACAGGAGCTGTGAATTCCTG 1721
Db 1572 TACGAAGAACAAAGTTACTTTTCGGTACGACAGCGGTTTACCGCGCGAACAATCGTAGCTG 1631
QY 1722 GACATCCATATCGATGTAAGTGTACGTTCCCAATGCGCGCGGCTGACCGATGTGA 1781
Db 1632 GGAAGATACCGAAATCAGAGTCAAGTTCCGAACGTGGCGCGCGCAACAGCGCGCTTAC 1691
QY 1782 AGTCACCGCGGTGGAGTTTCAGCAATCTGTATCTTACAAATTT-----TTGAGTGG 1835
Db 1692 GGTAAACGACGCGCGGCACTTACCAGCGAGCGTTTCAACAATTTTAACTGACTGACTGC 1751
QY 1836 AACGACAGATCGTTGTGTTTACTGTGAAAGTGGCTTCGACCAACCTGGGGATTA 1895
Db 1752 CGATCAGGTCACTGTCCGCTTCAAAGTCAAAATGCCACCGCGC---CTGGGACAAA 1808
QY 1896 GATTACTCAGCGGCAACATACCGAATTGGGAATTTGGACGCGATACGAGCGGAGC 1955
Db 1809 CGTCTACTGACCGGTAAAGTTCGCGAGCTTGGCACTGG-----ACAGCCGCCAA 1859
QY 1956 CGTTAAACAATGCGCAAGGCGCTGTGTCGCGCCCAATATCCGATGTGTTTATGTAT 2015
Db 1860 CGCAATCGGTCCGATGTACAACAGGTAGAAGCCAGTATCCGACTTGGTACTTCGACGT 1919
QY 2016 CAGGTTCCAGAGGAAGACCAATTCATCAAGTCTTCAATCAAGCGTGGATGGAAC 2075
Db 1920 CAGGTTTCCGGGCAACACCGCGCTGCAATTCAGTTTCATCAAGTGAACGGCTCGACGT 1979
QY 2076 GATTCAATGGGAGATGTTTCGAACCAAGTGGCCACAACCTCCACGGGTGCAACCGGTAA 2135
Db 1980 GACT---TGGGAGGCGGCAACACACCACTTCACTTCCGCTTCGACGCGGCTTGGAC 2036
QY 2136 CATTAATGTTTACGTGGCAAAAC 2157
Db 2037 CGTAACGCTCGATTGGCAGAAC 2058
```

```
RESULT 9
AAQ12366
ID AAQ12366 standard; cDNA; 2996 BP.
XX
AC AAQ12366;
XX
DT 25-MAR-2003 (updated)
DT 13-SEP-1991 (first entry)
XX
DE Gene encoding enzyme with starch degradation and cyclodextrin
DE synthesis activity.
XX
KW Crude starch; cyclodextrin; ss.
XX
OS Bacillus sp. B1018 (FERM P-9840).
XX
FH Key Location/Qualifiers
FT -35_signal 101..106
FT /*tag= a
FT -10_signal 125..130
FT /*tag= b
FT RBS 302..306
FT /*tag= c
FT sig_peptide 313..393
FT /*tag= d
FT mat_peptide 394..2451
FT /*tag= e
FT terminator 2567..2598
FT /*tag= f
FT /*note= "stem loop"
XX
PN JP03123491-A.
XX
PD 27-MAY-1991.
XX
PF 09-OCT-1989; 89JP-0262194.
XX
PR 09-OCT-1989; 89JP-0262194.
XX
PA (HGET ) HIGETA SHOYU KK.
XX
DR WPI; 1991-197926/27.
DR P-FSDB; AAR12743.
XX
PT Gene for enzyme which decomposes raw starch - and synthesises
PT cyclodextrin, DNA sequences from gene, vector and transformant
PS Claim 1; Fig 1; 9pp; Japanese.
XX
CC The gene can be ligated into an expression vector for mass prodn. of
CC the enzyme.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 2996 BP; 764 A; 856 C; 770 G; 606 T; 0 other;

Query Match 15.9%; Score 343.4; DB 12; Length 2996;
Best Local Similarity 51.8%; Pred.No. 9.3e-98;
Matches 1074; Conservative 0; Mismatches 921; Indels 78; Gaps 10;

QY 111 AAGCGTCAAAGGGACGTGATTTCACAGATTATCATTCACCGTTTTACGATGGGACAC 170
Db 423 AATTTCAGCACCGACGTCATCTATCAATTTTACCGACAGGTTTTTCGACGGCAATCC 482
QY 171 GAGCAACAAATCTCTGCCAAAAGTTATGACTTTAGATCCGACCAATTCGAATGGAA 230
Db 483 CGCCAAATTCGACCGC-----CGGCGGTTTACGGAACCTGCAGCACTCCG 533
QY 231 AATGTAATGGGCGGGATCTGGAGGGGTTTCGTCAAAACT-----TCCTTACTTAA 284
Db 534 GCTGTATTGGCGCGCGACATGGCAGGCCATCATCAACAAATCAACGAGGTTACTCTGAC 593
```

QY 285 ACAGTGGCGGTAAACGACAACTCTGTTGTCCTCCCGGTTTGGACAATCTGGATCACT--- 341
 Db 594 CGGATGGGGTTTACCGCCATCTGGATCTCCAGCCGGTGGAAACATCTACGATCAT 653
 QY 342 -----GGCGGGCAACGATAACACCGGGTATACAGGATATGAGACCGCGATTTAAACA 395
 Db 654 CAATTAATTCGGCGTCAACAACACCGCCATATCAAGGCTATGGGCCCGGATTTCAAGA 713
 QY 396 GATTGAGGAACATTTCGGGAAATTGGACACATTTGACACGTTGGTCAATGATGTCACCA 455
 Db 714 GACGATTCGGGCTACGGACCAATTTGCGGACTTTCAGAACCTTGATGCGCGCGGATGC 773
 QY 456 AAACGGATCAAGTGATTTGCGACTTTGTGCCCAATTCATTCGACATTCCTTTTAAAGCAAA 515
 Db 774 CAARAACATCAAAAGTCAATTAATGACTTCGCCCGCAACATACGTCGCCCGCTCGTCCGA 833
 QY 516 CGATTCCACCTTTCCGAGAGGGCGCCCTCTCAACAATGGAACCTATATGGCAATTA 575
 Db 834 CCAGCTTTCTTTTCGGGAAACCGGCCGCTGTACGATAACCGGACGCTGCTCGGGGATA 893
 QY 576 TTTTGATGACCAACAAAGGTTACTTCCACCATAATGGGACATCAGCAACTGGGACGA 635
 Db 894 CACGAACGATACGACGAACTGTTCACCATAC-----GGCGG 932
 QY 636 CCGGTACAGGCGCAATGGAATAAATTCACGGATCCAGCCGGTTTCTGCTTGGCGATTT 695
 Db 933 CACGGACTTTTCCACGACCGAAACCGCATCTACAAAACCTGTACGATCTCGCCGACCT 992
 QY 696 GTCCGAGGAAATGCGACGATGTCTCAATACCTCAGCCGATCGCGGGTTTCAATTTGGTAGC 755
 Db 993 GAACCATTAACAGACGCTCGGACGTCTATTGAGAGCGCATCAAAATGTGCTGGA 1052
 QY 756 ACATGACCGGATGGTTTGGATGATGCGGTGAAGCATTTTAAATCGGGGTTCTCAA 815
 Db 1053 TCTCGGCATCGACGCATCGCATGCGTGGTGAAGCATATGCCGTTTCGGGTGCGAGAA 1112
 QY 816 ATCGTTGCCGATAACTGTACCAAAAGAAAGACATTTTCTGTTGGGGAAATGTACGG 875
 Db 1113 GAGCTTTATGCTGCGCTGAACACTATAGCGCGTCTTTACCTTCGGGAAATGTTCT 1172
 QY 876 AGATGACCCCGGAACAGCCAAATCATCTGGAAAAGTTCGGTACGCCAACAAACAGCGGTG 935
 Db 1173 GGGGTAAACGAGTAGC-----CCGGAACCAATAGTTTGGCAACGATCCGCAT 1226
 QY 936 CAAATGCTGGAATTTGATCTCAACACGCTGATTCGAATGTGTTGCGCACATTTACGCA 995
 Db 1227 GAGCCTGTTGATTTCCGTTTCCCAAAAGTTCGGCGGTGTTTCCGGGCAACACCGA 1286
 QY 996 AACGATGTACGATCTTAACCAATATGTTGAACCAACGCGGAAACGAGTACAAATACAAAGA 1055
 Db 1287 CAAATGTAAGGCTGAGCGGATGCTGGAGGGCTCCGACCCGATTAACGCCAGTGA 1346
 QY 1056 AAATCTAATCAATTTATCGATAACCAATGATGTCAAGATTTCTTTTCGGTAAATGGA 1115
 Db 1347 TGACCAAGGTGACGTTCTATCGACCAACCATGACATGAGCGTTTCCACGCAAGCATGCAAA 1406
 QY 1116 CHAGGCAATTTGACAGCGGCTTCTTCAATCTCACTCTCGCGGTACCCCTCCAT 1175
 Db 1407 CGCCGGAAGTGGAGCAAGCGCTGCGGTTTCACTGATCTCGCGCGGTCCCGCCAT 1466
 QY 1176 CTATTATGAAACGAAACATACATGTCAGCGCGCAATGACCCGTACAAACCGGGGATGAT 1235
 Db 1467 TTATTACGCAACCGAGCAGTACATGTCGGTGGGACCGATCCGACACACCGGGCCGAT 1526
 QY 1236 CGCGGTTTGAATGACACACCGCTTTAAAGAGTGTCAACTCTCGCGGGTTTCG 1295
 Db 1527 CCCTTCTCTCCAGTTCGACACCGCCTATCAAGTCAATCAAAAGCTGCGCGCTCG 1586
 QY 1296 CAGGACAAATGGCGATCCAGTACGCAACCAACCCAGCGTTGGATCAACAATGATGT 1355
 Db 1587 CAAGTCAACCCGGCCATTCGCTACGATCGAGCGAGCGCTGGATCAACAACGAT 1646
 QY 1356 TTACATTTATGAACGGAATTTTCAACGATGTCGTTGGTGGCCATCAATCGAAACAC 1415

Db 1647 GCTCAATTTAGCGCAAAATTCGCGACGAACGTTCCCTCGTTGCCGTCAACGCAATTT 1706
 QY 1416 GCAATCCTCTATTCGATTTCCGTTTTCAGAGCGCTTTCGCAAAATGGCAGCTATCGGA 1475
 Db 1707 GAACGGCGCGCTTCCATTTCCGGAATTTGTCACCTTCCCTCCGCAAGCGAGCTACAAAGA 1766
 QY 1476 TTATCTGTCAAGGCTTGTGGGGGGGAACCGGATTTCCGTTTCCAAATGGAAGTGTGCTTC 1535
 Db 1767 CFTCTTGGCGCCCTTCTTGAAAGCGACACGTTAACGCTAGGCTCCGGGGAGCGCCCTC 1826
 QY 1536 G---TTACGCTTGGCCTGAGCCGCTGCTGTTGGCAGTACAGACATCCGCTTCAGC 1592
 Db 1827 CAATTTCAAGCTTGGCGCGCGGACCGCGGTGTGGCAGTACACCGCGGCAACGCGCAC 1886
 QY 1593 GCGCAAAATCGGATCGGTTGCTCCAAATATGGGATTCGCGGTAAATGTGTCACGATCGA 1652
 Db 1887 GCGACCATCGGCAATGTCGGGCGGATGATGGCCAGCGCGGCTGACGATCAOGATCGA 1946
 QY 1653 CGGAAAGTGTGGGACGACGAGGGAACCGTGACATTTGGCGGAGTGACAGCTG- 1711
 Db 1947 CGCGCGCGCTTCCGCTCTAGCAAAAGCACCGCTTACTTTCGGCACGACGCGGTGAGCGG 2006
 QY 1712 -----TGAAATCCTCGACATCCATCGGATGGAATGAGTGTACGTTCCCAACATGCGCGC 1763
 Db 2007 CGCAACATCAGCTTTGGGAAGACCGCGATCAAGTGAAATTCGCGCGCTCGCAG 2066
 QY 1764 CGGCTGACCGATGTGAAAGTFC---ACCGGGGTGGAGTTTCCAGCAATCTGTATTTT 1820
 Db 2067 CGGCATCTACAATTAAGTTCGCAACCGCGGAAACGCGCAACGCTGTACGACAA 2126
 QY 1821 CAATTTTGAAGTGAACGACGACATCGGTTGTTGTTTACTGTGAAAAGTGGCTCCGAC 1880
 Db 2127 CTTGAGGTATTTTCGGAGACCGAGTACGCTCGCTTCGTTGCTCAACACGCGCAAC 2186
 QY 1881 CAACCTGGGGGTAAGATTTTACCTGACGGCAACATACCGGAATTTGGGAATTTGGAGCAC 1940
 Db 2187 GGCCTTGGGCAAAATCTTACTGACGGCAATGTACGAGCTGGGAACTGGGACCC 2246
 QY 1941 GGAATGAGCGAGCGGCTTAAATGAGCGAAGGCGCCCTGCTCGCGCCCAATTTATCGGA 2000
 Db 2247 G-----GCAAAAGCGATCGGGCGGATGTACAACACAGTCTGTTACCAATATCCGAA 2297
 QY 2001 TTGCTTTTATGATTCAGGTTTCCAGCAGCAAGAGATTCATTCAGTTCTTCATCAA 2060
 Db 2298 CTGATATATGACGTACGCTTCCGCGCGCAAAAGATTCGAGTTCAAGTTTGAAGA 2357
 QY 2051 GCGTGGGATGGAACGATTCATATGGGAGATGTTTCGAACCAACCGTGGCCCAACTCCAC 2120
 Db 2358 AC---AAGGCTCCACCGTCACGTGGGAAGCGCGACGACCAACACCTTCACCGCGCGTC 2414
 QY 2121 GGTGCAACCGGTAAACATTTCTGTACGTGGA 2153
 Db 2415 CAGCGCACCGGACCATTAACGTGAATGGCA 2447

RESULT 10

AAQ50345
 ID AAQ50345 standard; DNA; 2414 BP.

AC AAQ50345;

XX 02-APR-1994 (first entry)

DE Wild type cyclomulodextrin glucanotransferase gene.

XX Alpha-CD; mutant; alpha-cyclodextrin; CGTase;
 KW starch; beta-CD; ss.

XX Bacillus stearothermophilus.

OS Key Location/Qualifiers
 FH CDS 213..2348
 FT

```

FT FT /*tag= a
FT -35_signal 9..14
FT /*tag= b
FT -10_signal 32..37
FT /*tag= c
FT repeat_region 105..137
FT /*tag= d
FT /note= "inverted repeat region"
FT RBS 197..205
FT /*tag= e
FT /note= "Shine-Dalgarno sequence"
FT repeat_region 2352..2395
FT /*tag= f
FT /note= "inverted repeat region"
PN JP05244945-A.
PD XX
PD 24-SEP-1993.
PF XX
PF 04-MAR-1992; 92JP-0082843.
PX XX
PX 04-MAR-1992; 92JP-0082843.
PR XX
PR (NISO ) NIPPON SHOKUIN KAKO KK.
PA
XX WPI; 1993-347473/44.
DR P-PSDB; AAR42844.
XX
XX The sequence is that of the wild type Bacillus stearothermophilus
CC cyclomultodextrin glucanotransferase (CGTase) gene. The gene may be
CC mutated by site directed mutagenesis to encode a protein having a
CC F222Y and/or a F286Y mutation. The substituted CGTase is mixed with
CC 8 percent starch and produces alpha-cyclodextrin instead of beta-CD.
XX
SQ Sequence 2414 BP; 732 A; 417 C; 570 G; 695 T; 0 other;

Query Match 15.2%; Score 328.6; DB 14; Length 2414;
Best Local Similarity 51.6%; Pred. No. 4.2e-93;
Matches 978; Conservative 0; Mismatches 869; Indels 48; Gaps 8;

QY 277 TATCTTAAACAGCTGGGCTAACGACAACTGCTGGTGTCCCGGTTTGGHCAATCTGGAT 336
DB 489 TATTTAACAGATATGGGTGTACAGCGATATGGGATTTCTCAGCCTGTAGAAAATGTATT 548
QY 337 ACACCTG-----GCGGCGACCGATAACACGGGCTATCAAGGATCTGACGCGCGATTT 390
DB 549 TCTGTGATGATGATGACAGCGGTTACGCTCTCTATCATGTTATTGGGCGCGGATTC 608
QY 391 AAACAGATTGAGAAACATTTCCGGAATTGGACACATTTGACACGTTGTCATATGCT 450
DB 609 AAAAAGCCAAACCGGTTTTTGGTACCGCTCAGTGATTTCCAAACGTTAGTTGATGCGCA 668
QY 451 CACCAAAACGGAATCAGGTGATTTGCTGACTTTGTGCCAATCAATCGACTCCCTTTAAG 510
DB 669 CATGCAAAAGGAATAAGGTATTTATGACTTTTGCCCCCAACCATATCTTCTCTGCTTCA 728
QY 511 GCAAAACGATTCCACCTTTTGGGAAGGCGGCGCTCTACAAATGGAACCTATATGGGC 570
DB 729 GAAACGAATCTCTTATATGAAACCGACGACTGTACGATATGGACATGCTTGGC 788
QY 571 AATTATTTTATGACGACAAACAAAGGGTACTTCCACCATATATGGGACATCAGCAACTGG 630
DB 789 GGTTCACAAATGATG--CCAAATGATTTTCCACATAACGG-----TGG 832
QY 631 GAACACGGGTACGAGCGCAATGGAACAACTTCAAGATTCAGCGGTTTCTCGCTTGGC 690
DB 833 AACAAACGTTTCCAGCTTAGAGATGGGATTTATCGAAATCTG-----TTTGACTTGGCG 887

```

```

QY 691 GATTTGCGCAGGAAATGGCAGATTGCTCAATACCTGACCCGATGCGCGGTCAATTG 750
DB 888 GACCTTAACCATCAGAAACCTGTTATTGATAGGTATTAAAAGATGCAGTAAAAATGG 947
QY 751 GTAGCACATGAGCGGATGTTTGGGATTGATGCGGTGAAGCATTTTAATTCGGGGTTC 810
DB 948 ATAGATATGGGATTGATGTTATCGGTATCGGTATGCGGTGAGCACATCCCGTTGGATG 1007
QY 811 TCCAAATCGTTGGCCGATAACTGTACAAAAGAAAGACATTTTCTCGTGGGGAATGG 870
DB 1008 CAAAAATCTCTGATGATGAGATTGATAACTATCGTCTGCTTTTACGTTTGGGAGTGG 1067
QY 871 TACGGAGATGACCCCGGAACACCACTCACTGGAAGAGTCCGCTACGCCAACACAGC 930
DB 1068 T-----TTTGTGCAAAAATGAAGTGGACGCGAAACATCATTTTCCCAATGAAGT 1121
QY 931 GGTGTCAATGTGCTGATTTTGTCTCAACACCGGTGATTGAAATGTTGTCGGCACATTT 990
DB 1122 GGAATGAGTTTGTCTCGATTTTGTTCGGACAAAAGCTTCGTAAGTATTGGCAATAAC 1181
QY 991 AGCGAAACGATGTACGATCTTAAACAAATATGTTGAACCAACCGGGAACGATGCAATAC 1050
DB 1182 AGCGATAATTGGTATGGCTTTAATCAATGATTCAGATACCGCATCAGCATATGACGAG 1241
QY 1051 AAAGAAATCTAATCACAATTTATCGATAACCATGATATGTCAGATTTTCTTCGTTAAAT 1110
DB 1242 GTTCTGATCAAGTAACATTCATAGACAAACATGATATGATCGGTTTATGATGACGGA 1301
QY 1111 TCGAACAGCGGAATTTGACCAAGCGCTTGTCTTCAATTTCTCACTTCGCGGGTACSCCC 1170
DB 1302 GGAGATCCGCGCAAGGTGGATATGGCACTTGTCTATTATTGACATCCGTGGGTACCG 1361
QY 1171 TCCATCTATTATGGAACCGAACAAATACATGTCAGCGCGGCAATGACCCGTACACCGGGG 1230
DB 1362 AATATTTACTGTGTACAGAGCAATACATGACCGGTACCGGATCCGGAACAACTGTAAG 1421
QY 1231 ATGATGCGCGGTTTGATACGACCAACCCCTTTTAAAGAGGTGTCAACTTTCGCGGGG 1290
DB 1422 ATGATGAGTTCATTCATTAATAAATACTCGCGGTATCAAGTGATTCAAAATCTATCTCT 1481
QY 1291 TTCCGACGGAACAAATCGGCGATTCAGTACCGGACACCAACCGCGGTTGGATCAACAAT 1350
DB 1482 CTCGACGAAACAAATCCGCGGTAGCTTATGTTGATACCGAACAGCGTTGGATCAATGGC 1541
QY 1351 GATGTTTACATTTATGAAACGAAATTTTTCACCATGCTGCTGTTGTTGGCCATCAATCGA 1410
DB 1542 GATGTTATGTTATGAGCGACAGTTTGGCAAGATGTTGTTAGTTGCCGTTATGCT 1601
QY 1411 AACACGCAATCTCTTATTTGATTTTCGGTTTTCAGACGCGCTTGCCTAAATGGCAGCTAT 1470
DB 1602 AGTTCAAGCAATATTTACTCGATTACTGGCTTATTTACAGCTTTACCGACGAGACATAT 1661
QY 1471 GCGGATTTATCTGACGGCTGTTTGGGGGGAACGGGATT---TCGTTTCCAAATGGAAT 1527
DB 1662 ACAGATACGTTGGCGCTCTTTTAGACGGAATAACAATCAAGTCGGTTCAAATGGATCA 1721
QY 1528 GTGCGTCTGTTCAAGCTTGGCGCTGAGCGCTGCTGTTTGGCAGTACAGACATCCGCT 1587
DB 1722 GTTATGATTTGACTTAGGACCGGGGAAGTCCGTTGTTAGGCAATACAGTGCACAGAA 1781
QY 1588 TCAGCGCGCAAAATCGGATCGGTGCTCCAAATATGGGATTCGGGTTAATGTTGTCAG 1647
DB 1782 AGCAGCGCAATATTGCTGTCATGTTGACCGGATGATGGGCAAGTCGGTTCATCAAGTAAC 1841
QY 1648 ATCGACGGGAAGGTTTGGGACGCGCAGGGAACCGTGACATTTGGCGGATGACGCG 1707
DB 1842 ATTGATGCGGAAGGATTTCGGAACAAATACGGGCACTGTGAAGTTCGGAACGACAGCTCC 1901
QY 1708 ACTGGAATCTCTGACATCCATCGGATTTGAAGTGTACGTTCCCAACTATGGCGCGGG 1767
DB 1902 AATGTTGCTCTGCTTACCAATCAATCGTTGCTGTACCAATGCTGTCCACGGA 1961
QY 1768 CTGACCGGATGTGAAGTTCACCGCGGCTGAGGTTTTCAGCAATCTGTATTTCTTACAATTT 1827

```

1962 AATATAATATTACCGTCCCAATCATCAAGCGTCAACAGAGTGGCGTTATGATAACTTT 2021
 1828 ----TTGAGTGAAGCGAGACATCGGTGTGTTTACTGTGAAAAGTGGCTCCGACCAAC 1884
 2022 GAAGTACTAACAAATGATCAAGTGTCAAGTGGGTTTGTGTTAATAACGCGACTACCAAT 2081
 1885 CTGGGGGTAAGATTACCTGACGGGCAACATACCGGAATGGGGAATTGGAGCACGGAT 1944
 2082 CTAGGCGCAATATATACATGTTGGCAACGATATAGCTCGGCACTGGGACACTAGT 2141
 1945 ACGACGGAGCGGTAAACAATGCGAAGGCGCCCTGCTGCGGCCCAATATCCGATGG 2004
 2142 AAG-----GCAATCGGTCCAAATGTTCAATCAAGTGGTTTACTCTCTATCTACATGG 2192
 2005 TTTTATGATTACAGGTTCCAGCAGGAAGAGGATTCAATTCAGTTCTTCATCAAGGT 2064
 2193 TATATAGATGTCAGTGTCCGAGGAAGAGACAATTGAGTTTAAATTATTAAGAAAGAC 2252
 2065 GCGGATGAAGCAATTCATATGGAGAAATGTTTCGAACACGCTGGCCCAACTCCACGGGT 2124
 2253 AGCCAGGTAAATGTCACCTTGGGAAAGTGGTTCAATCATGTTTATATACGACCAACGAAT 2312
 2125 GCAACGGTAACTACTGTATCGTGCAGAACTA 2159
 2313 ACACCGGAAAAATATAGTGGATTGGCAGAACTA 2347

RESULT 11

AAV28879 standard; DNA; 327 BP.

ID AAV28879 standard; DNA; 327 BP.
 AC AAV28879;
 XX
 DT 09-SEP-1998 (first entry)
 DE Nucleotide sequence of the starch binding domain E-domain.
 XX
 KW Starch binding domain; SBD; E-domain; oral; dental plaque;
 KW toothpaste; dental cream, mouthwash; ss.
 XX
 OS Bacillus steatothermophilus.

Key Location/Qualifiers

CDS 1..327
 /tag= a
 /product= "SBD E-domain"
 /note= "no start or stop codon specified"

W09816190-A1.

23-APR-1998.

13-OCT-1997; 97WO-DK00446.

11-OCT-1996; 96DK-0001128.

(NOVO) NOVO-NORDISK AS.

Fugisang CC, Tsuchiya R;

WPI; 1998-251040/22.

P-PSDB; AAW57249.

An oral care composition comprises a single unit Starch Binding
 PT Domain - useful for improving oral hygiene and preventing the
 PT formation and/or for removing dental plaque

Claim 11; Pages 19-20; 31pp; English.

This is the nucleotide sequence of the starch binding domain (SBD)
 CC E-domain used in the method of the invention to prepare an oral care
 CC composition. The product is useful for oral care, for

CC prevention/removal of dental plaque, as a toothpaste, dental cream,
 CC gel or tooth powder, odontic, mouth wash, pre- or post brushing rinse
 CC formulation, chewing gum, lozenge, or candy.

SQ Sequence 327 BP; 88 A; 77 C; 91 G; 71 T; 0 other;

Query Match 15.1%; Score 327; DB 19; Length 327;

Best Local Similarity 100.0%; Pred. No. 4.1e-93;

Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1831 AGTGAACGACGACATCGGTGTGTTTACTGTGAAAAGTGGCTCCGACCAACCTGGGG 1890
 DB 1 AGTGAACGACGACATCGGTGTGTTTACTGTGAAAAGTGGCTCCGACCAACCTGGGG 60

QY 1891 GATAAGATTTTACCTGACGCGGCAACATACCGGAATTGGGGAATTGGAGCAGGATACGAGC 1950
 DB 61 GATAAGATTTTACCTGACGCGGCAACATACCGGAATTGGGGAATTGGAGCAGGATACGAGC 120

QY 1951 GGAGCCGTTTAAACAATGCGCAAGGCCCCCTGCTGCGGCCCAATTATCCGGATGGTTTAT 2010
 DB 121 GGAGCCGTTTAAACAATGCGCAAGGCCCCCTGCTGCGGCCCAATTATCCGGATGGTTTAT 180

QY 2011 GTATTACGCGTTCACGACGAGAAAGACGATTCAATTCAAGTTCTTCATCAAGCGTGGGAT 2070
 DB 181 GTATTACGCGTTCACGACGAGAAAGACGATTCAATTCAAGTTCTTCATCAAGCGTGGGAT 240

QY 2071 GGAACGATTCAATGGGAGATGGTTGGAACACACGTCGCCACAACCTCCACAGGTTGCAACC 2130
 DB 241 GGAACGATTCAATGGGAGATGGTTGGAACACACGTCGCCACAACCTCCACAGGTTGCAACC 300

QY 2131 GGTAACTATTCTGTTACGTGGCMAAC 2157
 DB 301 GGTAACTATTCTGTTACGTGGCMAAC 327

RESULT 12

AAX23987
 ID AAX23987 standard; DNA; 327 BP.

AC AAX23987;
 XX
 DT 25-JUN-1999 (first entry)

DE B. steatothermophilus maltogenic amylase SBD E domain DNA.

XX Maltogenic amylase; starch-binding domain; SBD; E domain; DE domain.
 KW industry; proteolytic degradation; ss.
 XX
 OS Bacillus steatothermophilus.

W09915636-A1.

01-APR-1999.

23-SEP-1998; 98WO-DK00411.

07-OCT-1997; 97US-0061559.

24-SEP-1997; 97DK-0001100.

(NOVO) NOVO-NORDISK AS.

Andersen C, Frandsen TP, Jorgensen ST;
 WPI; 1999-254701/21.

P-PSDB; AAW93897.

Starch-binding domains from Bacillus steatothermophilus C599
 PT maltogenic amylase
 XX

Claim 2; Page 29-30; 36pp; English.

This invention describes the Bacillus steatothermophilus C599
 starch-binding domain (SBD), E and DE domains. The SBD coding sequences,

CC vectors and host cells are useful for the production of SBDs in a high
 CC yield. Extracellular production of the SBDs makes their use in
 CC industrial applications economically feasible. Also, the Bacillus host
 CC is capable of expressing a SBD without proteolytic degradation of the
 CC expressed protein or any problems normally associated with translocation
 CC of proteins in Bacillus hosts.

XX Sequence 327 BP; 88 A; 77 C; 91 G; 71 T; 0 other;

Query Match 15.1%; Score 327; DB 20; Length 327;
 Best Local Similarity 100.0%; Pred. No. 4.1e-93;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1831 AGTGAACACACATCGTTGTGTTTACTGTGAAGTGGCGCTCCGACCACTGGGG 1890
 DB 1 AGTGAACACACATCGTTGTGTTTACTGTGAAGTGGCGCTCCGACCACTGGGG 60
 QY 1991 GATPAGATTACTGACGGCAACATACCGGAATTGGGGAATTGGACGATACGAGC 1950
 DB 61 GATPAGATTACTGACGGCAACATACCGGAATTGGGGAATTGGACGATACGAGC 120
 QY 1951 GAGAGCGTTAACAATGCGCAGGCCCCCTGCTCGCGCCCAATATCCGGATTGGTTTAT 2010
 DB 121 GAGAGCGTTAACAATGCGCAGGCCCCCTGCTCGCGCCCAATATCCGGATTGGTTTAT 180
 QY 2011 GTATTGAGGTTCCAGCAGGAAGACGATTCAATTCAAGTCTTCATCAGCGTGGGAT 2070
 DB 181 GTATTGAGGTTCCAGCAGGAAGACGATTCAATTCAAGTCTTCATCAGCGTGGGAT 240
 QY 2071 GGAACGATTCAATTGGGGAATGTTGGAACCACTGGGCAACATCCCGGTTGCAACC 2130
 DB 241 GGAACGATTCAATTGGGGAATGTTGGAACCACTGGGCAACATCCCGGTTGCAACC 300
 QY 2131 GGTAACTTACTGTTACGTGCAAAAC 2157
 DB 301 GGTAACTTACTGTTACGTGCAAAAC 327

RESULT 13

AAQ01810
 ID AAQ01810 standard; DNA; 2565 BP.

XX AAQ01810;

XX 25-MAR-2003 (updated)

DT 10-MAR-2003 (updated)

DT 20-DEC-1990 (first entry)

XX Sequence encoding cyclomaltodextrin glucanotransferase (CGTase)
 DE having alkaline optimum pH.

XX Enzyme; ss.

XX Unidentified.

XX Key Location/Qualifiers
 FH 211..216
 FT RBS /*tag= a
 FT CDS 223..2142
 FT /*tag= b

XX JP02000455-A.

XX 05-JAN-1990.

XX 28-APR-1988; 88JP-0105945.

XX 28-APR-1988; 88JP-0105945.

XX (RIKA) RIKAGAKU KENKYUSHO.

XX WPI; 1990-047991/07.

DR P-PSDB; AAR06110.

XX Cyclo.malto:dextrin glucano:transferase coding DNA sequence -
 FT including corresp. plasmid and microbe transformed by plasmid
 XX Disclosure; Fig 5-1 to 5-2 Page 842; 16pp; Japanese.
 XX The DNA, the plasmid contg. it, and any microorganism transformed with
 CC plasmid are new. Transformed strains show CGTase activity after 16 hrs.
 CC culture on LB plate. 38-2 CGTase gene 17-1 CGTase genes are each stably
 CC retained in E.coli cells translated to CGTase efficiently. Large amts. of
 CC polypeptide with CGTase activity can be obtd.
 CC (Updated on 10-MAR-2003 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX

SQ Sequence 2565 BP; 663 A; 754 C; 629 G; 519 T; 0 other;

Query Match 15.1%; Score 326.2; DB 11; Length 2565;
 Best Local Similarity 52.1%; Pred. No. 2.5e-92;
 Matches 954; Conservative 0; Mismatches 823; Indels 54; Gaps 8;

QY 344 CGGGCACCGATAACACCGGGCTATCAGGATACGTGACGCGCATTTTAAACAGATTGAGG 403
 DB 572 CCGGGCTGAATAACACCGCATACCGGCTATTGGCTAGAGACTTCAAGAAAACGAACC 631
 QY 404 AACATTTCGGGAATTGACACACATTTGACACGTTGGTCAATGATGCTCACCAAAACGGA 463
 DB 632 CGGCTTACGGCACGATTGCGGATTTCCAGATTCATCGCTGCCGCCCATGCGCAAGACA 691
 QY 464 TCAAGTGATTGTGCGATTTGTGCCCAATCATTTGCTCCCTTTTAAAGGCAACGATTCCA 523
 DB 692 TCAAGTTCATCATGACTTCGACCGACACACAGTCTCCCGCTTCCTTGGACCGCGT 751
 QY 524 CTTTTGGGAAGGGCGGCCCTCTACACAAATGGAACCTATATGCGGCAATTTATTTGATG 583
 DB 752 CTTTTGGGAGACGCGCAAGCTGTACACACGCGCAGGACGAAGCGGATA---CAGGA 808
 QY 584 AGCAACAAAGGGTACTTCCACCATATGGGCACATCAGCACTGGGACGCGGTAGG 643
 DB 809 AAGTACGCACAATTTATTCATCATACGCGGGGACCGATTTTCGACGACCGAAGACG 868
 QY 644 AGCGCAATGGAAAAACTTCACGGATCCAGCGGTTTCTCGCTTGGCGATTTTGTGCGAGG 703
 DB 869 GGATTTCAAAA-----CCTGTAGATCTTGGGATCTGAACACA 910
 QY 704 AAATGGCACGATTGCTCAATACCTGACGATGCGGGGTTCAATTGGTAGCACTAGAG 763
 DB 911 ACACAGACACAGTGGATACCTACCTGAAGGATGCCATTAATAATGTGGCTCGATCTGGCA 970
 QY 764 CGGATGTTTGGGATGATGCGGTGAAGCATTTTAATTCGGGTTCTCCAATCGTTGG 823
 DB 971 TTGACGGCATTCGAATGACGCGGTAAACATATCCGTTCCGCTGGCAGACAGCTTCA 1030
 QY 824 CCGATAAACTGTACAAAAGAAAGACATTTTCTGTGGGGGAATGTTAGCGAGATGACC 883
 DB 1031 TGCGCACCGCTCAACAATTACACGCGGTTTCACTTCGCGAATGTTCTCTCGGGTAA 1090
 QY 884 CGGAAACAGCCAAATCATCTGAAAAGGTCCGGTAGCCACACACAGGGTGTCAATGTGC 943
 DB 1091 ACGAAGTGAG-----TGCGGAGATCATAGTTTCGCCAAGTTTCGCGCATGAGCTGC 1144
 QY 944 TGGATTTTGATCTCAACACCGGTGATTTCGAAATGTGTTGCGCACATTTTACGAAAACGATGT 1003
 DB 1145 TGGATTTTCCGCTTCGCCAGAGTTTCGCCAGTATTTCAGAGGACACACCGCAATATGT 1204
 QY 1004 ACGATTTTAAACAATATGTTGAACCAAAACGGGGAACGAGTACAAATACAAAGAAATCTAA 1063
 DB 1205 ACGGACTGAATCCATGCTAGAGGGCTCTGGACCGATTACGCGCAGATGGAGGATCAGG 1264
 QY 1064 TACATTTATCGATAACCATGATATGTCAGATTTCTTTTGGTAAATTCGAAACAGGCGA 1123
 DB 1265 TCACGTTTCATGATACACGACATGGAACGTTTCCACAATAACAGCGCAACACCGCCGA 1324

```
Qy 1124 ATTTCACCGGCGCTTGCTTTCAATCTCACTTCGCGGGGTACGCCCTCAATCTATTATG 1183
Db 1325 AGCTGGAGCAAGCGCTCGCTTACAGTTAACTTCCCGCGGTGCGAGCACTATTATG 1384
Qy 1184 GAACCGACATACATGCGAGCGGCATGACCGTACACCGGGGATGATGCGGCGT 1243
Db 1385 GAACCGACATACATGTCGGCGGGAACGATCCCGAACCGGGGCTCGAATCTCTTT 1444
Qy 1244 TTGATAGCAACACCGCGCTTTAAAGAGGTGTCAACTCTGGCGGGTGTGGCGAGAAC 1303
Db 1445 TCTCTACACGACTACGGCTTATCAGTCACTAAAGCTGGCGCTCTGGTAATCCA 1504
Qy 1304 ATGCGCGATCCAGTACGGACACACCCAGCGTTGGATCAACAAATGATTTACATTT 1363
Db 1505 ATCCGGCATCGCTTATGGAGCAGACCGAGGAGCGCTGGATAACACAGATGTGTGATCT 1564
Qy 1364 ATGAACGGAATTTTCAACGATGCTGTGTGGGCCATCATATGAAACACGCAATCCT 1423
Db 1565 ATGAGCGAGTTGGGCAACACGATAGCCGTCATCGCGTCAACCGCAAGTGAAACAGT 1524
Qy 1424 CATTATCGATTCGGTTTCAGACGGCGCTTGCAGATGGCAATGGACGTATGATCTGT 1483
Db 1625 CGGCTCATTACGGACTGTAACTCACTGCGCGCGGACGTACACCGACGTGCTGG 1684
Qy 1484 CAGGCTGTTGGGGGGAAGGATTCGGTTTCCATGGAATGCGCTTC--GTCA 1540
Db 1685 GCGGCTCTTGAACGCTAACATTTGACCGTGGCTCGCGGATCCGCTTCCATCTTCA 1744
Qy 1541 CGCTTGCGCTTGGAGCGCTGTCTGTTGGCAGTACAGACATCCGCTTCAGCGCGCAAA 1600
Db 1745 CGCTGGCAGCGGCGGTACGGCGCTGTGCAATACACACCGCGGTGACGGCCACCA 1804
Qy 1601 TCGATGCGTTGTCCTCAATATGCGGATTCGGGATTCGGGTATATGTCAGATCGACGGGAAG 1660
Db 1805 TCGGCAGTAGGACGATGATGGCCAAACACGAGCGAGCGCTCACCAATGACGCGCGG 1864
Qy 1661 GTTTTGGAGCAGCAGGGAACCTGACATTTGGCGAGTGACA-----GCGACTG 1711
Db 1865 GTTTCGGTGACAAAGGACAGGTATCTTCGGCAGCAGCGGAGTTACGGGCGCTAATA 1924
Qy 1712 TGAATTCCTGGACATCCAAATCGGATGAAGTGTACGTTCCCAACATGGCGCGCGGTGA 1771
Db 1925 TTACGGCTTGGGAAGACACGACATCAAAAGTGAATATCCCTGCGCTTGGCGAGCGTAT 1984
Qy 1772 CGATGTCAAGTCAACGCGGT---GGAGTTCCAGCATCTGTATTTTACATATTT 1828
Db 1985 ACAATATCAAAATCGCCACAGTGGCGGAACCTCAAGCAATGTGCATGACAACTTCGAAG 2044
Qy 1829 TGAGTGAACGCAGACATCGGTTGTGTTTACTGTGAAAGTGGCGCTCCGACCAACCTGG 1888
Db 2045 TGCTGAGCGGGATCAGTACAGCTGCGCTTTGTGTGAACACGCAACACCGCGCTCG 2104
Qy 1889 GGGATAGATTACCTGACGGGCAACATACCGGAATTTGGGAATTTGGAGACGATACGA 1948
Db 2105 GCCAGACGTATATCTGTCAGGACGCGTAAGCAGCTCGGCAACTGGGACCGGCCA-- 2161
Qy 1949 GCGGAGCGTTTAAATGCGCAAGGCGCCCTGCTCGCGCCCAATTTACGGAATGTTTT 2008
Db 2162 -----AAGCCTCGCTCGCTGTGACACAGGTCATCTACCAATACCAACCTGGTACT 2215
Qy 2009 ATGTATTCAGCGTTCCAGCAGGAAGACGATTCAATTTCAAGTTCTTTCAATCAAGGTGCGG 2068
Db 2216 ATGACGTCAACGTTCCGCGCGCAAAACGATCGAATTTAAATTCCTGAAAAACA--GG 2272
Qy 2069 ATGGACGATTCATATGGAGAAATGTTTGAACACGATGGGCAACATCTCCACGGGTGAA 2128
Db 2273 GCTCAGCGGTACGTTGGGAAGCGGCTCGAACCATACTTCACCGCAGCAACGAGCGGCA 2332
Qy 2129 CCGGTAACATTAATCTGTTACGTTGGCAAACTA 2159
Db 2333 CGGCCACCAATAACGTGAATGCAACCGTA 2363
```

```
RESULT 14
AAQ10164
ID AAQ10164 standard; DNA; 2659 BP.
XX
AC AAQ10164;
XX
DT 19-MAR-1991 (first entry)
XX
Cyclic maltodextrin glucotransferase 17-1 gene.
XX
CGTase; ss.
XX
Bacillus strain 17-1 (ATCC 31007).
XX
Key Location/Qualifiers
FT CDS 224..2365
FT /tag= a
FT /product= CGTase 17-1
FT 217..722
FT RBS
FT /tag= b
FT sig_peptide 224..304
FT /tag= c
FT mat_peptide 305..2362
FT /tag= d
FT misc_RNA 224..1003
FT /tag= e
FT misc_RNA 1004..1540
FT /tag= f
FT /label= U1
FT /tag= U2
FT misc_RNA 1541..1831
FT /tag= g
FT /label= U3
FT misc_RNA 1832..2362
FT /tag= h
FT /label= U4
XX
JP02286080-A.
XX
26-NOV-1990.
XX
28-APR-1989; 89JP-0110603.
XX
28-APR-1989; 89JP-0110603.
XX (RIKA ) RIKAGAKU KENKYUSHO.
XX
WPI; 1991-012216/02.
DR P-PSDB; AAR10052.
XX
New cyclomalto:dextrin glucanotransferase - has DNA sequence
PT coding amino acid sequence
XX
Disclosure; Fig 6; 16pp; Japanese.
XX
The sequence is taken from pUCP1. The sequence is divided into
CC four units (U1-U4) by restriction sites. By combining these with
CC similar units (S1-S4) from CGTase 38-2 (AAQ10163), e.g. S1-U2-U3-S4
CC or U1-S2-U3-S4, enzymes with properties of both strains can be
CC produced. The hybrids compensate for defects in the individual
CC enzymes and can be designed to display optimum characteristics of
CC pH and temperature stability.
CC See also AAQ10163.
XX
SQ Sequence 2659 BP; 694 A; 782 C; 653 G; 530 T; 0 other;
```

```
Query Match
Best Local Similarity 15.0%; Score 324.6; DB 12; Length 2659;
Matches 953; Conservative 0; Mismatches 824; Indels 54; Gaps 8;
Qy 344 CGGCACCGGATACACCGGGCTATCACGATCTGGACGCGCATTTTAAACAGATTGAGG 403
Db 573 CCGGCGTGAATAACACCGCATACCGCTATTGGCTTAGAGCTTCAAGAAACGACCC 632
```

QY 404 AACATTTGGGAATTTGGACACCAATTTGACACGTTGGTCAATGATGTCACCAAAACGAA 463
DB 633 CGGCTTACGGCAGGATTCGGAATTTCCAGATCTCATCGCTGCGCCATGCCAGAAC 692
QY 464 TCAAGTGATTTGTCGATTTGTCGCAATCAATTCGATCTCTTTTAAGGCAACGATTTCA 523
DB 693 TCAAGTGATTCATGATCTTCGCAACCGAACCAACGATCTCTCTGCTTTTGAACGAGCGT 752
QY 524 CTTTTCGGAAGGCGGCGCTCTCAACAATGGAACCTATATGCGCAATTTATTTGATG 583
DB 753 CTTTTCGGAAGGCGGCACTGTACACAACGCGCAGGACGAGCGGATTA---CAGCA 809
QY 584 AGCAACAAAGGATTTCTTCAACATATGCGGCAATTCAGCACTGGAGCAGCGGTACG 643
DB 810 ACGATACGCAAAATTTATTCATATACGCGGGGACCGATTTCTGACGACCGGAGACG 869
QY 644 AGCGCAATGGAACAACTTCAACGATCCAGCGGTTCTGCTTGGCGATTTGTCGAGG 703
DB 870 GATTTCAAAA-----CCTGTAGATCTTGGCGATCTGAACCA 911
QY 704 AATAGGCAAGATTTGCTCAATACCTGACCGATGCGGGGTTCAATTTGGTAGACATGGAG 763
DB 912 ACAACAGCAGATGATACCTACTGAGGATGCCATTAATAATGCTGCTCGATCTGGCA 971
QY 764 CGGATGTTTGGGATTTGTCGCTGAGCATTTTAATTCGGGTTCTCCAAATCGTTGG 823
DB 972 TTGACGGATTCGAATGACGCGGTAAACATATGCGCTTCGGCTGGCAGAGATTTCA 1031
QY 824 CCGATAACTGTACCAAAAGAAAGACATTTCTCTGGGGAAATGTCGAGAGATGACC 883
DB 1032 TGSCCAACCTGACAAATTAACGCGGTCTTCACTTCGGCGAATGTTCTCTCGGGTAA 1091
QY 884 CGGACACGCAATCTGTAAGAGTCCGGTACGCAACACAGCGGTGTCATATGTC 943
DB 1092 ACGAAGTGAG-----TGCGGAGATCATAAAGTTTCGCCAACGTTTCGGCATGAGCCTGC 1145
QY 944 TGAATTTGATCTCAACACGATGATTCGAAATGTTGTCGGCACATTTAGCGAAACGATGT 1003
DB 1146 TGAATTTGCTTCCGCGAAGGTTTCGCGAGTATTCAGGACACACGCAATATGT 1205
QY 1004 ACGATCTTAAACATATGTTGAAACCAACCGGGGAGAGTACAATACAAAGAAATCTAA 1063
DB 1206 ACGGACTGAAATCCATGCTAGAGGCTCTGACGCGATTAACGCGAGATGAGGATCAG 1265
QY 1064 TCAATTTATCATACCATGATATGTCAGATTTCTTCGGTAAATTCGAACAGGCGA 1123
DB 1266 TCAAGTTTCATGATAACCAACAGATGGAACGTTTCCAAATAACAGCGCCACCGCCGA 1325
QY 1124 ATTTGACACGAGCGCTTCTTTCAATCTCACTTCGCGGGGTACGCCCTCCATCTATTATG 1183
DB 1326 AGCTGGAGCAAGCGCTCGCTTCAAGTTAACTTCGCGGGGTGCGAGCAATCTATTATG 1385
QY 1184 GAACCGAACAATATGATGACGAGGCGAATGACCGGTACACCGGGGATGATGCGGGGT 1243
DB 1386 GAACCGGACAATATGATGTCGCGGGGAGAGATCCGACACCGGGCTTCGAATTCCTTT 1445
QY 1244 TTGATACGACCAACCGCTTTTAAAGAGTGTCAACTCTGCGGGGTTGCGCAGGAACA 1303
DB 1446 TCTCTACGAGCTACGCTTATCGGTCAGTAAAGTGTGCGCCCTCTCGGTAAATCA 1505
QY 1304 ATGCGGATTCAGTACGCGACCCACCGAGGTTGGATCAACAATGATGTTTCAATTT 1363
DB 1506 ATCCGCGCATTCGCTTATGGGACGACGCGAGGCGCTGGATAAACAAGATGTTGATCT 1565
QY 1364 ATGAACGGAATTTTCAACGATGTCGTGTTGGTGGCCATCAATCGGAAACAGCATCT 1423
DB 1566 ATGAGCGCAAGTTCGCGAACAACGATGACGCTCATCGCGTCAACCGCAACGTAACAGT 1625
QY 1424 CTTATTCGATTTCCGCTTTCAGACGCGCTTTCGCAAAATGGCAGCTATGCGGATTCGT 1483
DB 1626 CGGCTTCATACGGGATGTTGATACCTCACTGCGCGCGGAGCTACACCGAGTCTGTGG 1685

QY 1484 CAGGCTGTGTGGGGGGGAAACGGGATTTCCGTTTCCATATGGAAGTGTGCTTC---GTTCA 1540
DB 1686 CGGCTCTTTTGAACGGTAACAAATTTGACCTCGCTCGCGGGATCGCTTCCATCTTCA 1745
QY 1541 CGCTTGGCTCGAGCGCTGTCTGTTGGCAGTACAGCACATCGCTTTCAGCGCGCAAA 1600
DB 1746 CGCTGGCAGCGCGCGGTACGGCGCTGTGGCAATACACACGCGCGTGACTGCGGCAACCA 1805
QY 1601 TCGGATCGGTGTCTCAAAATATGGGATTTCCGGGTAATGTGTCAAGTACGATCGAGGGAAG 1660
DB 1806 TCGGCGACGTAGGACCGGATGATGCCAAACAGCGCGACGCTGACCATTTGACGCGCGG 1865
QY 1661 GTTTTGGACGACGAGGAAACCGTGACATTTTGGCGGAGTGACA-----GCCACTG 1711
DB 1866 GTTTTGGTGAACCAAAAGGCAACGGTATCTTTCGGCAGCAGCGCATTTACCGGGGCTAAATA 1925
QY 1712 TGAATCTCGGATCTCAATCGGATTCGAAGTGTACGTTTCCCAACATGCGCGCGGCTGA 1771
DB 1926 TTACGGCTTGGGAGACACGCGAGATCAAAGTGAATTCCTGCGCTTGGCGAGGCGTAT 1985
QY 1772 CCGATGTGAAGTACACCGGGGT---GGAGTTTCCAGCAATCTGATTTCTTACATATTT 1828
DB 1986 ACAATATCAAAATCGCAACAGTTCGCGAACTCAAGCAATGTGCATGACAACTTCGAAG 2045
QY 1829 TGAGTGAACGCGACATCGGTGTTGTTTACTGTGAAAGTGGCTCCGACCAACTGG 1888
DB 2046 TGCTGACGGGATCAGTCAAGCTGCGCTTGTGTGAAACACGACACCGCGCTG 2105
QY 1889 GGGATAAGATTTACCTGACGCGCAACATACCGGAATTTGGGAATTTGGAGCAGGATAGA 1948
DB 2106 GCCAGAACGTATATCTGCGAGCAGCTAAGCAGAGCTCGGCAACTGGGACCGGCA--- 2162
QY 1949 CGGAGCGGTTAACATGCGCAAGGCGCTGCTCGCGCCCAATTCGCGATTTGTTTT 2008
DB 2163 -----AAGCATCGCTGCTGTGTAACAGGTCTTACCAATCCCACTTGTACT 2216
QY 2009 ATGATTCAGCGTTCCAGCAGGAAGACGATTCAAATTCAGTTCTTCATCAAGCTGCGG 2068
DB 2217 ATGACGTACCGTTCCGCGCGCAACGATCGAATTTAAATTCCTGAAACAA---GG 2273
QY 2069 ATGGAACGATTCATGCGGAATGTTGCAACACGCGGCGCAACTCCACGCGGTGCA 2128
DB 2274 GCTGACGGTAACGTGGGAAGCGCTCGAACCATCTTACCGCACCAACCGAGCGCA 2333
QY 2129 CGGTAAACATTTACTGTTACGTGSCAAACTA 2159
DB 2334 CGGCCACATAAACGTGAACCTGCAACCGTA 2364

RESULT 15
AAT85105
ID AAT85105 standard; DNA; 2040 BP.
XX
AC AAT85105;
XX
DT 25-MAR-2003 (updated)
DT 19-NOV-1997 (first entry)
XX
XX Cyclomaltodextrin glucanotransferase activity protein encoding DNA.
DE Food; drink; glycosyl transfer; microbial; starch; taste; ss.
XX
XX Bacillus stearothermophilus.
XX
XX Key Location/Qualifiers
FT CDS 1..2040
FT /*tag= a
FT /product= Cyclomaltodextrin glucanotransferase
FT activity protein
FT /note= "No stop codon given"
XX
XX JP09187289-A.

PD 22-JUL-1997.
 XX 30-MAR-1986; 97JP-0025738.
 XX 30-MAR-1986; 86JP-0072290.
 PR 30-MAR-1986; 86JP-0025738.
 XX (HAYB) HAYASHIBARA SEIUTSU KAGAKU.
 XX WPI; 1997-419402/39.
 DR P-PSDB; AAW27108.
 XX Preparation of a food/drink product - by using glycosyl transfer
 PT reaction
 XX Example 4; Page 14; 16pp; Japanese.
 XX A method has been developed for the preparation of a food/drink product.
 CC The method involves reacting a starch product with a microbially-derived
 CC Bacillus stearotheophilus polypeptide. The polypeptide has the
 CC following properties: (i) has molecular weight 70 plus or minus 10 kD;
 CC (ii) thermostable to 70 degree C or greater; (iii) has cyclomaltodextrin
 CC glucanotransferase activity; and (iv) contains a specifically claimed
 CC amino acid sequence in the N-terminal portion of the protein. The
 CC present sequence encodes the Bacillus stearotheophilus
 CC cyclomaltodextrin glucanotransferase activity protein. The food/drink
 CC product is safe and has good taste, proper texture and a long shelf
 CC life.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX Sequence 2040 BP; 621 A; 359 C; 486 G; 574 T; 0 other;
 SQ
 Query Match 14.9%; Score 321.8; DB 18; Length 2040;
 Best Local Similarity 51.4%; Pred. No. 5.5e-91;
 Matches 973; Conservative 0; Mismatches 872; Indels 48; Gaps 8;
 QY 277 TATCTTAACAGCTGGCGTAACGACAACTGCTGTGTCCTCGGTTTGGACAACTCGAT 336
 DB 184 TATTTAAACAGATATGGGTGTAACGGATATGGATTTCTCAGCCTGTAGAAATGATTT 243
 QY 337 ACACTG-----GGGGCACCGGATAACACGGGCTATCACGGATCTAGGACGGCGGATTTT 390
 DB 244 TCTGTGATGAATGATGCAAGCGTTCCGCATCCTATCATGTTATTGGCGCGGATTTTC 303
 QY 391 AAACAGATGAGGACAACTTCGGGAATGGACCACATTTGACAGCTGTGTCATGATGCT 450
 DB 304 AAAACGAAACCCGTTTTTGTACCTCTAGTGAATTCACAGTTTCCACGTTTGTATGATCCGCA 363
 QY 451 CACCAAAACGGAATCAAGGTGATTCGACTTTGTGCCCAATCATTCGACTCCTTTTAAAG 510
 DB 364 CATGCAAAAGGAATAAGGTAATTTGACTTTGCCCAACCATACTTCTCTGCTTCA 423
 QY 511 GCAACAGATTCCACTTTGCGGAAGCGCGCCCTCTCAACAAATGGAACCTTATATGGGC 570
 DB 424 GAAACGAATCTCTTATATGAAACGGAACGACTGTACGATATGGACATTTGCTTGGC 483
 QY 571 AATTTATTTGATGACGCAACAAAGGTACTTCCACCATAATGGGACATCAGACACTGG 630
 DB 484 GGTTCACAAATGATG---CCACATGTTATTTTACACATAACG-----TGG 527
 QY 631 GACGACCGGTACGAGGCGCAATGGAATACTTCACGATCCAGCGGTTTCTCGCTTGCC 690
 DB 528 AACACGTTTTCCAGCTTAGAGGATGGATTTATCGAAATCTG-----TTTGACTTGGC 582
 QY 691 GATTTGTCGAGGAATAATGGACGATTTGCTCAATACCTGACCGATGGCGGTTCAATG 750
 DB 583 GACCTTAACCAATCAGAACCTTGTATTGATAGGTTATTTAAAGATGCAGTAAAAATGTGG 642
 QY 751 GTAGCACATGGAGCGGATTTGCGGATTTGATGGGTGAAGCATTTTAAATTCGGGGTTC 810
 DB 643 ATGATATGGGATTTGATGTTATCCGTTATGGAATGGGTAAGCAATCATGCGGTTTGATGG 702
 QY 811 TCCAAATCGTTGGCGGATAAACTGTATCCAAAAAGAACATTTTCTCTGGTGGGGAATGG 870

DB 703 CAAAATCTCTGATGATGAGATTGATAAATATATCGTCCTGTCTTTTACGTTTGGGAGTGG 762
 QY 871 TACGGAGATGACCCCGGAACAGCCCAATCATCTGGAAAGGTCCGGTAGCCACACACAGC 930
 DB 763 T-----TTTGTGAGAAATGAAAGTGGAGCGGAAACAATCATTTACTTTGGCAATGAAAGT 816
 QY 931 GGTGTCAATGTGCTGATTTTGTGATCTCAACACGGTGAATTCGAAATGTGTTGGGCACATTT 990
 DB 817 GGAATGAGTTTCTCGATTTTTCGTTTCGGACAAAAGCTTCGTCAAGTATTTCGCGAATAAC 876
 QY 991 ACGCAACAGATGTAGATCTTAAACAATPATGTGTACCAAAACGGGGAACAGGTACAAATAC 1050
 DB 877 AGCGATATTTGGTATGGCTTTTAAATCAAAATGATTACAGATACGGCATCAGCATATGACGAG 936
 QY 1051 AAGAAAAATCTAATCACATTTATCGATAACATGATGTCAGATTTCTTCGGTAAAT 1110
 DB 937 GTTCTCGATCAAGTAAACATTCATAGAACCAATGATATGGATTCGGTTTATGATGACGGA 996
 QY 1111 TCGAACAAAGCGGAATTTGACACAGGCGCTTTGCTTTTCAATTTCTCACTTCGCGGGGTACGCC 1170
 DB 997 GGAGATCCGCGCAAGGTGATATGGCACTTGTCTGTATTATTGACATCCCGTGGCTACCG 1056
 QY 1171 TCCATCTATTATGGAAACGACAAATACATGCGAGCGGCAATGACCCGTACACCCGGGG 1230
 DB 1057 AATATTTTACTATGTACAGAGCAATACATGACGGTATACGGGATCCAAACAATTCGTAAG 1116
 QY 1231 ATGATGCGCGCTTTGATACGACAAACACCGCTTTTAAAGAGGTGTCAACTCTGCGGGGG 1290
 DB 1117 ATGATGAGTTCAATCAATAAAATACTTCGCGGTATCAAGTGAATCAAAACTATCTTCT 1176
 QY 1291 TTGCGGAGGAACAAATGCGGGCATTCAGTAGGCAACACACCCAGCGTTGGATCAACAAT 1350
 DB 1177 CTCGACGAACAAATCCGCGTTAGCTTATGTTGATACGGAACAGCGTTGGATCAATGGC 1236
 QY 1351 GATGTTTACATTTTATGACGGAATTTTCAAGATGCTGTGTGTTGGTGGCCATCAATCGA 1410
 DB 1237 GATGTGATGTGATAGGACAGTTTGGCAAGATGTTGTGTAGTTTCGGGTTAATCGT 1296
 QY 1411 AACACGCAATCTCTCTATTTCGATTTCCGGTTTCAGACGGCTTGCACAAATGGACGTAT 1470
 DB 1297 AGTTCAAGCAGTAATTTACTCGATTACTGGCTTATTTTACAGCTTTACAGACGGAACATAT 1356
 QY 1471 GCGGATTTATCTGTCAGGCTGTGGGGGGAACGGATTT---TCCGTTTCCAATGGAAGT 1527
 DB 1357 ACGGATCAGCTTGGCGGTCTTTTAGACGGAATACAAATTCAAAGTCGTTTCAATGGATCA 1416
 QY 1528 GTGCTTTCGTTCAAGCTTCGCGCTGGAGCGGTCTGTGTTTGGCAGTACAGCAATCCGCT 1587
 DB 1417 GTTAATGCAATTTGACTTAGGACCGGGGAAGTCCGTGTATGGCATACAGTGCAACGAA 1476
 QY 1588 TCAGCGCGCAATCGATCGGTTGCTCAAAATPATGGGATTCGGGTAATGTGGTCAAG 1647
 DB 1477 AGCACGCAATTTATGTCATGTTGGAACCGATGATGGGCAAGTCGGTCATCAAGTAAAC 1536
 QY 1648 ATCAACGGAAGAGTTTGGGACGACGACGAGCAACGTCACATTTGGCGAGTGAACGCG 1707
 DB 1537 ATTATGCGAGGAATTCGGAACAAATACCGGCACATGTGAAGTTCGGAACGACAGTGGC 1596
 QY 1708 ACTGTGAAATCTCTGGAATCCAAATCGGATTTGAAAGTGTAGCTTTCCCAACATGCGCGCGGG 1767
 DB 1597 AATGTTGTGTGTCATAACAAATCAAAATCGTTGTGGCTGTACCAATGTGTACACAGGA 1656
 QY 1768 CTGACCGATGTGAAGTACACCGGGTGGAGTTTCCAGCAATCTGTATTCTTACAAATTT 1827
 DB 1657 AAATATAATATTACCGTCCAAATCATCAAGCGGTCAAAACGAGTGGCGCTTATGATACTTT 1716
 QY 1828 ---TTGAGTGGAAACGACAGCATTCGGTTGTGTTTACTGTGTAAGGTGGCGCTCCGACCAAC 1884
 DB 1717 GAAGTACTAACAATGATCAGTGTGAGTGGGTTTGTGTTTAAATACGCACTACCAAT 1776
 QY 1885 CTGGGGATTAAGATTTTACCTGACCGGCAACATACCGGAATTTGGGGAATTTGAGCAAGGAT 1944

Db 1777 CTAGGCGCAAAATATATATACATGTTGTGGCAACGTATATAGAGCTCGGCACTGGGACACTAGT 1836
QY 1945 ACGAGCGGAGCGGTTTAAACAATGCGCAAGGGCCCCCTGCTCGCGCCCAATTATCCGGATTGG 2004
Db 1837 AAG-----GCAATCGGTCCAATGTTCAATCAAGTGGTTTACTCCTATCCTACATGG 1887
QY 2005 TTTTATGTAATTCAGCGTTCCAGCAGGAAAGACGATTCATTCAGTTCTTCATCAGCGT 2064
Db 1888 TATATAGATGTCAGTGTCCCGAAGGAAAGACAATTGAGTTTAAAGTTTATTAAGAAAGAC 1947
QY 2065 GCGGATGGAAACGANTTCATGCGGAGATGGTTCGAACCACTGGGCCAACCTCCACGGGT 2124
Db 1948 AGCCAAAGGTAAATGTCACTTGGGAAAGTGGTTCAATCATGTTTATAGCACCAACGAT 2007
QY 2125 GCACCGGTAAACATTACTGTGTACGTGGCAAAAC 2157
Db 2008 ACAACCGGAAAAATTATAGTGGATTGGCAGAAC 2040

Search completed: September 17, 2003, 08:30:39
Job time : 431 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 08:17:27 ; Search time 112 Seconds
(without alignments)
8512.389 Million cell updates/sec

Title: US-10-069-908-1
Perfect score: 2160
Sequence: 1 atgaaagaaagcgttc.....ctgtacgtggcaaaactag 2160

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/FcTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2160	100.0	2160	3	US-09-386-607-1 Sequence 1, Appli
2	2160	100.0	2160	4	US-09-645-707B-1 Sequence 1, Appli
3	576	26.7	576	3	US-09-376-071-3 Sequence 3, Appli
4	327	15.1	327	3	US-09-376-071-1 Sequence 1, Appli
5	318.2	14.7	2133	3	US-08-947-965-1 Sequence 1, Appli
6	296	13.7	2061	1	US-08-204-656B-9 Sequence 9, Appli
7	296	13.7	2061	1	US-08-470-702-5 Sequence 5, Appli
8	296	13.7	2061	1	US-08-467-831-5 Sequence 5, Appli
9	231.4	10.7	2582	2	US-08-816-105A-2 Sequence 2, Appli
10	208.2	9.6	2100	1	US-07-927-316A-1 Sequence 1, Appli
11	74	3.4	1988	1	US-08-469-202-26 Sequence 26, Appli
12	74	3.4	1988	2	US-08-484-434C-33 Sequence 33, Appli
13	74	3.4	1988	4	US-09-384-361-33 Sequence 33, Appli
14	74	3.4	2000	1	US-08-469-202-25 Sequence 25, Appli
15	74	3.4	2000	2	US-08-484-434C-32 Sequence 32, Appli
16	74	3.4	2000	4	US-09-384-361-32 Sequence 32, Appli
17	45	2.1	509	4	US-09-634-238-173 Sequence 173, App
18	42.2	2.0	7218	1	US-08-232-463-14 Sequence 14, Appli
19	42	1.9	4411529	3	US-09-103-840A-1 Sequence 1, Appli
20	40.4	1.9	403765	3	US-09-103-840A-2 Sequence 2, Appli
21	39.6	1.8	672	4	US-09-634-238-208 Sequence 208, App
22	39.4	1.8	1404	1	US-08-204-656B-1 Sequence 1, Appli
23	39.4	1.8	1404	1	US-08-204-656B-7 Sequence 7, Appli
24	39.4	1.8	1404	1	US-08-470-702-1 Sequence 1, Appli
25	39.4	1.8	1404	1	US-08-470-702-4 Sequence 4, Appli
26	39.4	1.8	1404	1	US-08-467-831-1 Sequence 1, Appli
27	39.4	1.8	1404	1	US-08-467-831-4 Sequence 4, Appli

28 39 1.8 1650 4 US-09-107-532A-3186 Sequence 3186, Ap
29 37.8 1.7 1404 1 US-08-204-656B-3 Sequence 3, Appli
30 37.8 1.7 1404 1 US-08-204-656B-5 Sequence 5, Appli
31 37.8 1.7 1404 1 US-08-470-702-2 Sequence 2, Appli
32 37.8 1.7 1404 1 US-08-470-702-3 Sequence 3, Appli
33 37.8 1.7 1404 1 US-08-467-831-2 Sequence 2, Appli
34 37.8 1.7 1404 1 US-08-467-831-3 Sequence 3, Appli
35 37.4 1.7 2489 4 US-09-596-824-1 Sequence 1, Appli
c 36 37.4 1.7 2489 4 US-09-885-329-1 Sequence 1, Appli
37 37 1.7 6142 4 US-09-514-302-1 Sequence 1, Appli
38 35.8 1.7 3600 1 US-08-537-002A-5 Sequence 5, Appli
39 35.8 1.7 3600 3 US-08-863-010-5 Sequence 5, Appli
40 35.8 1.7 3600 3 US-09-024-429-5 Sequence 5, Appli
41 35.6 1.6 500 3 US-09-141-000-2 Sequence 2, Appli
42 35.4 1.6 2415 3 US-09-308-922-1 Sequence 1, Appli
43 35 1.6 882 6 5338678-5 Patent No. 5338678
44 35 1.6 1692 4 US-09-107-532A-2582 Sequence 2582, Ap
45 35 1.6 1692 4 US-09-107-532A-2583 Sequence 2583, Ap

ALIGNMENTS

RESULT 1
US-09-386-607-1
; Sequence 1, Application US/09386607
; Patent No. 6162628
; GENERAL INFORMATION:
; APPLICANT: Cherry, Joel
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Beier, Lars
; APPLICANT: Evansen, Torben
; TITLE OF INVENTION: Maltogenic Alpha-Amylase Variants
; FILE REFERENCE: 5443.414-US
; CURRENT APPLICATION NUMBER: US/09/386,607
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: DK98/00269
; EARLIER FILING DATE: 1998-02-27
; EARLIER APPLICATION NUMBER: 60/077,795
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2160)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (100)..(2157)
US-09-386-607-1

Query Match 100.0%; Score 2160; DB 3; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAAGAAACGCTTCTTTATTTGCGGACTGATGCTCTCATCGGTCTTCTGTTTC 60
DB 1 ATGAAAAAGAAACGCTTCTTTATTTGCGGACTGATGCTCTCATCGGTCTTCTGTTTC 60
QY 61 ACGGCTTCTCTCGGTACATCCAAACCGCGCTGACGACGAGTTCGCCAGCGTCAAA 120
DB 61 ACGGCTTCTCTCGGTACATCCAAACCGCGCTGACGACGAGTTCGCCAGCGTCAAA 120
QY 121 GGGGACGTGATTTACGAGATTATTCATTGACCGGTTTTACGATGGGACACGACGACAAAC 180
DB 121 GGGGACGTGATTTACGAGATTATTCATTGACCGGTTTTACGATGGGACACGACGACAAAC 180
QY 181 AATCTCGCAAAAGTTATGACTTTTACGATCCGACCAAAATCGAAGTGGAAAAATGTATTGG 240

Db	181	AATCCTCCCAAAGTATTGACATTTCAGATCCGACCAAAATCGAAGTGGGAAAATGATTGG	240
Qy	241	GGCGGGATCTGGAGGGGTTTCGTCAAAAACCTTCCTTATCTTAAACAGCTGGCGGTAAACG	300
Db	241	GGCGGGATCTGGAGGGGTTTCGTCAAAAACCTTCCTTATCTTAAACAGCTGGCGGTAAACG	300
Qy	301	ACAAATCTGGTTCTCCCGGTTTTCGGCAATCTCTGGATACTCTGGCGGGCACCGATACACG	360
Db	301	ACAAATCTGGTTCTCCCGGTTTTCGGCAATCTCTGGATACTCTGGCGGGCACCGATACACG	360
Qy	361	GGCTATCACGGATACTGGACGCGCGATTTTAAACAAGATTGAGGAAACAATTTTCGGGAATTTGG	420
Db	361	GGCTATCACGGATACTGGACGCGCGATTTTAAACAAGATTGAGGAAACAATTTTCGGGAATTTGG	420
Qy	421	ACCAATCTGACACAGTTGGTCTCAATGATCTCACCAAAAACCGAATCAAGTGATTGTGCAC	480
Db	421	ACCAATCTGACACAGTTGGTCTCAATGATCTCACCAAAAACCGAATCAAGTGATTGTGCAC	480
Qy	481	TTTGTGCCCAATCATTCAGACTCCTTTTAAAGGCAAAACGATCCACCTTTTGGGAAAGCGCGC	540
Db	481	TTTGTGCCCAATCATTCAGACTCCTTTTAAAGGCAAAACGATCCACCTTTTGGGAAAGCGCGC	540
Qy	541	GCCTCTTACAAATAGBAACCTATATAGGCAATTAATTTTGATGACGCAACAAAAGGTAC	600
Db	541	GCCTCTTACAAATAGBAACCTATATAGGCAATTAATTTTGATGACGCAACAAAAGGTAC	600
Qy	601	TTCCACCATTAATGGGACATCAGCAACTGGGACGACCGGTACGAGGCGCAATGGGAAAAC	660
Db	601	TTCCACCATTAATGGGACATCAGCAACTGGGACGACCGGTACGAGGCGCAATGGGAAAAC	660
Qy	661	TTCCACCGAATCCAGCGGTTTCCTGCTTCGCCGATTTGTTCGACGAAAATGGCACGATTGCT	720
Db	661	TTCCACCGAATCCAGCGGTTTCCTGCTTCGCCGATTTGTTCGACGAAAATGGCACGATTGCT	720
Qy	721	CAATACCTGACGATGCGCGGTTCAATTTGATGACACATGAGGCGGATGTTTCGCGATT	780
Db	721	CAATACCTGACGATGCGCGGTTCAATTTGATGACACATGAGGCGGATGTTTCGCGATT	780
Qy	781	GATCGGTGAGCAATTTAAATTCGGGGTTCTCCAAATCGTTTGGCGGATAAACTGTACCAA	840
Db	781	GATCGGTGAGCAATTTAAATTCGGGGTTCTCCAAATCGTTTGGCGGATAAACTGTACCAA	840
Qy	841	AAGAAAGACATTTTCTGCTGGGGATGTGACGAGATGACCCCGGAACCGCAATCAT	900
Db	841	AAGAAAGACATTTTCTGCTGGGGATGTGACGAGATGACCCCGGAACCGCAATCAT	900
Qy	901	CTGGAAGAGTCCGGTAGCCCAACAAACAGCGGTGTCAATGTCTGTGATTTTCAATCTCAAC	960
Db	901	CTGGAAGAGTCCGGTAGCCCAACAAACAGCGGTGTCAATGTCTGTGATTTTCAATCTCAAC	960
Qy	961	ACGCTGATTCGAATGTGTTTCGGACATTTTACGAAACGATGTACATCTTACCAATATG	1020
Db	961	ACGCTGATTCGAATGTGTTTCGGACATTTTACGAAACGATGTACATCTTACCAATATG	1020
Qy	1021	GTGAAACCAACCGGGAAACGAGTACAAATACAAAGAAATCTAATCACATTTATCGATAAC	1080
Db	1021	GTGAAACCAACCGGGAAACGAGTACAAATACAAAGAAATCTAATCACATTTATCGATAAC	1080
Qy	1081	CATGATATGCAAGATTTCTTTCGGTAAATTCGAAACAGGCGAATTTGCACCGGCGGTT	1140
Db	1081	CATGATATGCAAGATTTCTTTCGGTAAATTCGAAACAGGCGAATTTGCACCGGCGGTT	1140
Qy	1141	GCTTTTCATTTCTCATTTTCGGGGGTACGCCCTCCATCTATTATGAAACCGAACAATACATG	1200
Db	1141	GCTTTTCATTTCTCATTTTCGGGGGTACGCCCTCCATCTATTATGAAACCGAACAATACATG	1200
Qy	1201	GACGCGGCAATCACCCGTACAAACCGGGGATGATCCCGGGTTTGTATACGAAACCCACC	1260
Db	1201	GACGCGGCAATGACCCGTACAAACCGGGGATGATCCCGGGTTTGTATACGAAACCCACC	1260
Qy	1261	GCCTTTAAGAGTGTCAACTCTGCGGGGTTTCGCGAGGAACAATTCGCGGATTCAGTAC	1320
Db	1261	GCCTTTAAGAGTGTCAACTCTGCGGGGTTTCGCGAGGAACAATTCGCGGATTCAGTAC	1320

Qy	1321	GGCACACCAACCCACGCTTGGATCAACAATGATGTTACATTTATGAAACGGAAATTTTTC	1380
Db	1321	GGCACACCAACCCACGCTTGGATCAACAATGATGTTACATTTATGAAACGGAAATTTTTC	1380
Qy	1381	AACGATGTGCTGTGGTGGCCATCAATCCGAACACGCAATCCCTCTATTTCGATTTCCGCT	1440
Db	1381	AACGATGTGCTGTGGTGGCCATCAATCCGAACACGCAATCCCTCTATTTCGATTTCCGCT	1440
Qy	1441	TTGCAGACGGCCTTCCCAAAATGGCAGCTATGCGGATTTCTGTACGGGCTGTGGGGGGG	1500
Db	1441	TTGCAGACGGCCTTCCCAAAATGGCAGCTATGCGGATTTCTGTACGGGCTGTGGGGGGG	1500
Qy	1501	AACGGGATTTCCGTTTCCAAATGGAGTGTGCTTCAAGCTTCGGCTTCGGACCGCGTG	1560
Db	1501	AACGGGATTTCCGTTTCCAAATGGAGTGTGCTTCAAGCTTCGGCTTCGGACCGCGTG	1560
Qy	1561	TCTGTTTGGCAGTCACGACATCCGCTTCAGCGCGCGAAATCGATTCGGTTCGTCCTCAAAAT	1620
Db	1561	TCTGTTTGGCAGTCACGACATCCGCTTCAGCGCGCGAAATCGATTCGGTTCGTCCTCAAAAT	1620
Qy	1621	ATGGGGAATCCGGGTAATGTGGTCAACGATCGACGGGAAAGGTTTTCGGACGACGCGAGGA	1680
Db	1621	ATGGGGAATCCGGGTAATGTGGTCAACGATCGACGGGAAAGGTTTTCGGACGACGCGAGGA	1680
Qy	1681	ACCGTGACATTTGGCGGAGTCACAGCGATGTGAATCCCTGGACATCCCAATCGGATTTGAA	1740
Db	1681	ACCGTGACATTTGGCGGAGTCACAGCGATGTGAATCCCTGGACATCCCAATCGGATTTGAA	1740
Qy	1741	GTGTACGTTTCCCAACAATGGCCGCGCGGCTGACCGCATGTGAAGTCAACCGCGGGTGGAGTT	1800
Db	1741	GTGTACGTTTCCCAACAATGGCCGCGCGGCTGACCGCATGTGAAGTCAACCGCGGGTGGAGTT	1800
Qy	1801	TCCAGCAATCTGTATTTCTTACATATTTTGAATGGAACGACGACATCGGTTGTGTTTACT	1860
Db	1801	TCCAGCAATCTGTATTTCTTACATATTTTGAATGGAACGACGACATCGGTTGTGTTTACT	1860
Qy	1861	GTGAAAAGTGGCCTCCGACCAACCTTGGGGATAGATTTTACCTGACGGGCAACATACCG	1920
Db	1861	GTGAAAAGTGGCCTCCGACCAACCTTGGGGATAGATTTTACCTGACGGGCAACATACCG	1920
Qy	1921	GAATGGGGAAATGGAGCACGATACGACGGAGCCGTTTAACTACGCAAGGGGCCCTG	1980
Db	1921	GAATGGGGAAATGGAGCACGATACGACGGAGCCGTTTAACTACGCAAGGGGCCCTG	1980
Qy	1981	CTCGGCCCAATATTCGCGATTTGGTTTATGTATTTACGCTTCCACGACGAAGACGATTT	2040
Db	1981	CTCGGCCCAATATTCGCGATTTGGTTTATGTATTTACGCTTCCACGACGAAGACGATTT	2040
Qy	2041	CAATTTCAAGTCTTTCATCAACGCTTGGGATGGAAACGATTCAAATGGGAGATGTTTCGAAC	2100
Db	2041	CAATTTCAAGTCTTTCATCAACGCTTGGGATGGAAACGATTCAAATGGGAGATGTTTCGAAC	2100
Qy	2101	CACGTGSCACAACTCCCAACGGGTGMAACGGGTAACTACTGTATCGTGGCAAACTAG	2160
Db	2101	CACGTGSCACAACTCCCAACGGGTGMAACGGGTAACTACTGTATCGTGGCAAACTAG	2160

RESULT 2
US-09-645-707B-1
; Sequence 1, Application US/09645707B
; Patent No. 6482622
; GENERAL INFORMATION:
; APPLICANT: Cherry, Joel
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Beier, Lars
; APPLICANT: Frandsen, Torben
; APPLICANT: Schafer, Thomas
; TITLE OF INVENTION: Amyolytic Enzyme Variants
; FILE REFERENCE: 5241.204-US
; CURRENT APPLICATION NUMBERS: US/09/645,707B
; CURRENT FILING DATE: 2000-08-24

```

; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2157)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (100)..()
; OTHER INFORMATION:
US-09-645-707B-1

Query Match      100.0%; Score 2160; DB 4; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGAAAACGGCTTTCTTTATTTGGGACCTGATGCTCTCATCGGTCTTCTGTTTC 60
DB 1 ATGAAAAAGAAAACGGCTTTCTTTATTTGGGACCTGATGCTCTCATCGGTCTTCTGTTTC 60
QY 61 AGCGGTTCTCTTCGGTACAAATCCAAACGCGCTGAAGCCAGCAGTTCCGCAAGCGTCAAA 120
DB 61 AGCGGTTCTCTTCGGTACAAATCCAAACGCGCTGAAGCCAGCAGTTCCGCAAGCGTCAAA 120
QY 121 GGGGACGTGATTACAGATTATCATTGACCGGTTTACGATGGGACACGACGAAAC 180
DB 121 GGGGACGTGATTACAGATTATCATTGACCGGTTTACGATGGGACACGACGAAAC 180
QY 181 AATCTCTGCCAAAAGTTATGACCTTACGATCCGACCAAAATCGAAGTGGAAATGATTGG 240
DB 181 AATCTCTGCCAAAAGTTATGACCTTACGATCCGACCAAAATCGAAGTGGAAATGATTGG 240
QY 241 GCGGGGATCTGGAGGGGGTTCGTCAAAAATCTCTTATCTTAAACAGCTGGGCGTAAAC 300
DB 241 GCGGGGATCTGGAGGGGGTTCGTCAAAAATCTCTTATCTTAAACAGCTGGGCGTAAAC 300
QY 301 ACAATCTGGTTGTCGCCGGTTTGGACAACTCGAATACATCTGGCGGGGACCGAATAAC 360
DB 301 ACAATCTGGTTGTCGCCGGTTTGGACAACTCGAATACATCTGGCGGGGACCGAATAAC 360
QY 361 GCGTATCAGGATCTGACGCGGATTTTAAACAGATTGAGAAACATTTCCGGGAATGG 420
DB 361 GCGTATCAGGATCTGACGCGGATTTTAAACAGATTGAGAAACATTTCCGGGAATGG 420
QY 421 ACCACATTTGACAGTTGGTCAATGATGCTACCAAAACGGAATCAAGGTGATTCGAC 480
DB 421 ACCACATTTGACAGTTGGTCAATGATGCTACCAAAACGGAATCAAGGTGATTCGAC 480
QY 481 TTTGTGCCCAATCTCGACTCTCTTTAAGGCAACAGATTCCACCTTTCCGGAAGGGCG 540
DB 481 TTTGTGCCCAATCTCGACTCTCTTTAAGGCAACAGATTCCACCTTTCCGGAAGGGCG 540
QY 541 GCGCTCTACAACATGAACTATATGSGCAATTTTGTATGACGACCAACAAAGGGTAC 600
DB 541 GCGCTCTACAACATGAACTATATGSGCAATTTTGTATGACGACCAACAAAGGGTAC 600
QY 601 TTCCACCAATAATGGGACATCAGCAACTGGGACGACCGGTACGAGGCGCAATGGAAAAAC 660
DB 601 TTCCACCAATAATGGGACATCAGCAACTGGGACGACCGGTACGAGGCGCAATGGAAAAAC 660
QY 661 TTCACGGAATCCAGCGGGTTCTCGCTTTCGGATTTCGACGAGAAATGGACGATTGCT 720
DB 661 TTCACGGAATCCAGCGGGTTCTCGCTTTCGGATTTCGACGAGAAATGGACGATTGCT 720
QY 721 CAATACCTGACCGGATCGGGGTTCAATTGTTAGCACATGACGCGGATGGTTGCGGATT 780
DB 721 CAATACCTGACCGGATCGGGGTTCAATTGTTAGCACATGACGCGGATGGTTGCGGATT 780
QY 781 GATCGGTTGAAGCATTTTAAATTCGGGGTTCCTCCAAATCGTTGCGCGATAACTGTACCA 840
DB 781 GATCGGTTGAAGCATTTTAAATTCGGGGTTCCTCCAAATCGTTGCGCGATAACTGTACCA 840
QY 841 AAGAAGACATTTTCTGTTGGGGAAATGTTACGAGATGACCCCGGAAACAGCCAATCAT 900
DB 841 AAGAAGACATTTTCTGTTGGGGAAATGTTACGAGATGACCCCGGAAACAGCCAATCAT 900
QY 901 CTGGAAGAGTCCGGTACGCCAACACAGCGGTGCTAATGTGTGATTTTTCATCTCAAC 960
DB 901 CTGGAAGAGTCCGGTACGCCAACACAGCGGTGCTAATGTGTGATTTTTCATCTCAAC 960
QY 961 ACGGTGATTCGAAATGTTTCGSCACATTTACGCAACAGGATGTACGATCTTACCAATATG 1020
DB 961 ACGGTGATTCGAAATGTTTCGSCACATTTACGCAACAGGATGTACGATCTTACCAATATG 1020
QY 1021 GTGAACCAACCGGGAAACAGATACAAATACAAAGAAATCTAATCACTTTATCGATAAC 1080
DB 1021 GTGAACCAACCGGGAAACAGATACAAATACAAAGAAATCTAATCACTTTATCGATAAC 1080
QY 1081 CATGATATGCAAGATTTCTTTCCGTAAATTCGAACAGGCGAATTTGCACAGCGGCTT 1140
DB 1081 CATGATATGCAAGATTTCTTTCCGTAAATTCGAACAGGCGAATTTGCACAGCGGCTT 1140
QY 1141 GCTTTCAATTCCTACCTTCGCGGGGTACGCCCTCCATCTATTATGSAACCGAAACAATACATG 1200
DB 1141 GCTTTCAATTCCTACCTTCGCGGGGTACGCCCTCCATCTATTATGSAACCGAAACAATACATG 1200
QY 1201 GAGCGGCAATGACCGGTACAAACCGGGGGATGATGCGGGGTTTATACGACCAACACC 1260
DB 1201 GAGCGGCAATGACCGGTACAAACCGGGGGATGATGCGGGGTTTATACGACCAACACC 1260
QY 1261 GCTTTAAAGAGGTGTCACCTCTGCGGGGTTGCGAGGAAACAAATCGCGGATCCAGTAC 1320
DB 1261 GCTTTAAAGAGGTGTCACCTCTGCGGGGTTGCGAGGAAACAAATCGCGGATCCAGTAC 1320
QY 1321 GGCACCAACCCAGCGTTGGATCAACAATGATTTTACATTTATGAAACGGAATTTTTC 1380
DB 1321 GGCACCAACCCAGCGTTGGATCAACAATGATTTTACATTTATGAAACGGAATTTTTC 1380
QY 1381 AACGATGCTGTTGGTGGCCATCAATCGAAACAGCAATCTCTTATTCGATTTCCGGT 1440
DB 1381 AACGATGCTGTTGGTGGCCATCAATCGAAACAGCAATCTCTTATTCGATTTCCGGT 1440
QY 1441 TTGACAGCGGCTTGCCAATGSCAGCTATCGGATTTATCTGTCAGGCTGTGGGGGG 1500
DB 1441 TTGACAGCGGCTTGCCAATGSCAGCTATCGGATTTATCTGTCAGGCTGTGGGGGG 1500
QY 1501 AACGGGATTTCCGTTTCCAAATGAGTGTGCTTCGTTACGCTTCGCGCTCGAGCCGTG 1560
DB 1501 AACGGGATTTCCGTTTCCAAATGAGTGTGCTTCGTTACGCTTCGCGCTCGAGCCGTG 1560
QY 1561 TCTGTTGGCAGTACAGCACAATCGCGTTACGCGCGCAAAATCGGATCGGTCTCCAAAT 1620
DB 1561 TCTGTTGGCAGTACAGCACAATCGCGTTACGCGCGCAAAATCGGATCGGTCTCCAAAT 1620
QY 1621 ATGCGGATTCGGGTAAATGTTGTCGATCGACGCGGAAAGGTTTGGGACGACGAGGGA 1680
DB 1621 ATGCGGATTCGGGTAAATGTTGTCGATCGACGCGGAAAGGTTTGGGACGACGAGGGA 1680
QY 1681 ACCGTGACATTTGGCGGAGTGAACGAGCTGTGAAATCTCGACATCCCAATCGGATTGAA 1740
DB 1681 ACCGTGACATTTGGCGGAGTGAACGAGCTGTGAAATCTCGACATCCCAATCGGATTGAA 1740
QY 1741 GTGTACGTTCCCAACATGCGCGCGGTGTGCCGATGTGAAATGTCAACCGCGGTGAGTT 1800
DB 1741 GTGTACGTTCCCAACATGCGCGCGGTGTGCCGATGTGAAATGTCAACCGCGGTGAGTT 1800
QY 1801 TCCAGCAATCTGATTTTACAAATTTTGTGAGTGAACGACATCCGTTGTTGTTACT 1860
DB 1801 TCCAGCAATCTGATTTTACAAATTTTGTGAGTGAACGACATCCGTTGTTGTTACT 1860
QY 1861 GTGAAAGTGGCGCTCCGACCAACCTGGGGGATGAAGATTTTACCTGACGGGCAACATCCG 1920
DB 1861 GTGAAAGTGGCGCTCCGACCAACCTGGGGGATGAAGATTTTACCTGACGGGCAACATCCG 1920
```

```
Db 1861 GTGAAAGTGGCGCTCCGACCAACCTGGGGGATAAGATTACTTGACGGGCAACATACCG 1920
Qy 1921 GAATTGGGGAATTGGAGACGGATACGACGGAGCGGTTAAACATGCGACAGGCGCCCTG 1980
Db 1921 GAATTGGGGAATTGGAGACGGATACGACGGAGCGGTTAAACATGCGACAGGCGCCCTG 1980
Qy 1981 CTCGCGCCCAATATCCCGAATTGGTTTATGTATTACGCGTTCCAGAGGAAGACGATT 2040
Db 1981 CTCGCGCCCAATATCCCGAATTGGTTTATGTATTACGCGTTCCAGAGGAAGACGATT 2040
Qy 2041 CAATTCAAGTTCTTCATCAGCGTGGGATGAGCGATTCATGCGGAGATGTTCCGAC 2100
Db 2041 CAATTCAAGTTCTTCATCAGCGTGGGATGAGCGATTCATGCGGAGATGTTCCGAC 2100
Qy 2101 CACGTGGCCACAACCTCCACGGGTGCAACCGGTAAACATTACTGTTACGTGGCAAACTAG 2160
Db 2101 CACGTGGCCACAACCTCCACGGGTGCAACCGGTAAACATTACTGTTACGTGGCAAACTAG 2160
RESULT 3
US-09-276-071-3
; Sequence 3, Application US/09276071
; Patent No. 6207149
; GENERAL INFORMATION:
; APPLICANT: Fugisang, Claus Crone
; APPLICANT: Tsuchiya, Rie
; TITLE OF INVENTION: Starch Binding Domains (SBDs) For Oral Care Products
; FILE REFERENCE: 5017.204-US
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Steatothermophilus
US-09-276-071-3
Query Match 26.7%; Score 576; DB 3; Length 576;
Best Local Similarity 100.0%; Pred. No. 5.8e-178;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1582 TCCTGCTTACGCGCCCAATCCGATCGTTGCTCCAAATATGGGATTCGGGTAAATGTG 1641
Db 1 TCCTGCTTACGCGCCCAATCCGATCGTTGCTCCAAATATGGGATTCGGGTAAATGTG 60
Qy 1642 GTCAAGTACGCGGAAGTTTGGACGACGCGAAGCCGTGACATTTGGCGGAGTG 1701
Db 61 GTCAAGTACGCGGAAGTTTGGACGACGCGAAGCCGTGACATTTGGCGGAGTG 120
Qy 1702 ACAGCGACTGTGAATCTCGACATCCATCGGATTTGAAGTACGTTCCCAACATGGCC 1761
Db 121 ACAGCGACTGTGAATCTCGACATCCATCGGATTTGAAGTACGTTCCCAACATGGCC 180
Qy 1762 GCGGGCTGACCGATGTGAAGTACCGCGGGTGGAGTTCCAGCAATCTGTATTCTTAC 1821
Db 181 GCGGGCTGACCGATGTGAAGTACCGCGGGTGGAGTTCCAGCAATCTGTATTCTTAC 240
Qy 1822 AATATTTGAGTGGAAACCGACATCGTTGTGTGTTTACTGTGAAAGTGGCGCTCGACC 1881
Db 241 AATATTTGAGTGGAAACCGACATCGTTGTGTGTTTACTGTGAAAGTGGCGCTCGACC 300
Qy 1882 AACTGGGGATAAGATTACTCAGCGGCAACATACCGAATTGGGAAATTCGAGCAG 1941
Db 301 AACTGGGGATAAGATTACTCAGCGGCAACATACCGAATTGGGAAATTCGAGCAG 360
Qy 1942 GATACGAGCGGAGCGGTTTAAACATGCGAAGGCGCCCTGCTCGGCGCCCAATATCCGAT 2001
Db 361 GATACGAGCGGAGCGGTTTAAACATGCGAAGGCGCCCTGCTCGGCGCCCAATATCCGAT 420
Qy 2002 TGGTTTATGTATTACGCGTTCCAGAGGAAAGACGATTCATTTCAAGTTCCTCATCAG 2061
Db 421 TGGTTTATGTATTACGCGTTCCAGAGGAAAGACGATTCATTTCAAGTTCCTCATCAG 480
```

```
Qy 2062 COTGCGGATGGAACGATTCAATGGGAGATGTTGGAACACCACTGGCCCACTCCCAAG 2121
Db 481 COTGCGGATGGAACGATTCAATGGGAGATGTTGGAACACCACTGGCCCACTCCCAAG 540
Qy 2122 GGTGCAACCGGTAAACATTTACTGTTTACGTGGCAAAAC 2157
Db 541 GGTGCAACCGGTAAACATTTACTGTTTACGTGGCAAAAC 576
RESULT 4
US-09-276-071-1
; Sequence 1, Application US/09276071
; Patent No. 6207149
; GENERAL INFORMATION:
; APPLICANT: Fugisang, Claus Crone
; APPLICANT: Tsuchiya, Rie
; TITLE OF INVENTION: Starch Binding Domains (SBDs) For Oral Care Products
; FILE REFERENCE: 5017.204-US
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Steatothermophilus
US-09-276-071-1
Query Match 15.1%; Score 327; DB 3; Length 327;
Best Local Similarity 100.0%; Pred. No. 1e-96;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1831 AGTGAACGAGACATCGGTTGTGTTACTGTGAAAGTGGCGCTCCGACCACTGGG 1890
Db 1 AGTGAACGAGACATCGGTTGTGTTACTGTGAAAGTGGCGCTCCGACCACTGGG 60
Qy 1891 GATAGATTTACTGACGGGCAACATACCGGAATGGGAAATGGGACGAGATACGAGC 1950
Db 61 GATAGATTTACTGACGGGCAACATACCGGAATGGGAAATGGGACGAGATACGAGC 120
Qy 1951 GGAGCGGTTAAACATGCGCAAGGGCCCTGCTCGCGCCCAATTTATCCGATTTGTTTAT 2010
Db 121 GGAGCGGTTAAACATGCGCAAGGGCCCTGCTCGCGCCCAATTTATCCGATTTGTTTAT 180
Qy 2011 GTATTACGCGTTCCAGAGGAAGACGATTCATTAAGTTTCTCATCAAGGTGCGGAT 2070
Db 181 GTATTACGCGTTCCAGAGGAAGACGATTCATTAAGTTTCTCATCAAGGTGCGGAT 240
Qy 2071 GGAACGATTCATGCGGAGATGTTGGAACCACTGGCCCACTCCCACTCCCGGTGCAAC 2130
Db 241 GGAACGATTCATGCGGAGATGTTGGAACCACTGGCCCACTCCCACTCCCGGTGCAAC 300
Qy 2131 GGTAAACATTTACTGTTACGTGGCAAAAC 2157
Db 301 GGTAAACATTTACTGTTACGTGGCAAAAC 327
RESULT 5
US-08-947-965-1
; Sequence 1, Application US/08947965A
; Patent No. 6004790
; GENERAL INFORMATION:
; APPLICANT: Dijkhuizen, Lubbert
; APPLICANT: Dijkstra, Hauke
; APPLICANT: Andersen, Carsten
; APPLICANT: Osten, Claus von der
; TITLE OF INVENTION: Cyclomaltodextrin Glucanotransferase
; FILE REFERENCE: 4285.204-US
; CURRENT FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 0477/95
```

; EARLIER FILING DATE: 1995-04-21
 ; EARLIER APPLICATION NUMBER: 1173/95
 ; EARLIER FILING DATE: 1995-10-17
 ; EARLIER APPLICATION NUMBER: 1281/95
 ; EARLIER FILING DATE: 1995-11-16
 ; EARLIER APPLICATION NUMBER: PCT/DK96/00179
 ; EARLIER FILING DATE: 1996-04-22
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 2133
 ; TYPE: DNA
 ; ORGANISM: Thermoanaerobacter sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (82)...(2130)
 US-08-947-965-1

Query Match 14.7%; Score 318.2; DB 3; Length 2133;
 Best Local Similarity 51.9%; Pred. No. 2.6e-93;
 Matches 1133; Conservative 0; Mismatches 958; Indels 90; Gaps 15;

QY 7 AAGAAAACGCTTTCTTTATTTGGGACTGATGCTCTCTCATCGTCTTCTGTTTCAGCGGT 66
 DB |||||
 4 AAGAAAACGCTTTAAACTTCTCTCGATTCTGTTGATAACCAATGCTCTCTTTTCAGC--- 60
 QY 67 TCTCTTCCTGACAAATCCAAACGCGCTGAAGCCAGCAGTTCGGCAAGCTCAA-----A 120
 DB |||||
 61 TCAATTCCTAATCCGTCGCGGAGCAGCAGTCTCAGTTTCCAAATGTTGCAATTAATCA 120
 QY 121 GGGGAGCTGATTACAGATTATCATTTACCGGTTTTACGATGGGGACAGCAGCAACAAC 180
 DB |||||
 121 ACAGATGTAATCTACAGATAGTCAAGACCGTTTTTTAGATGGATCCAGTAATAAT 180
 QY 181 ATTCCTGCCAAAAGTTATGGACTTTAGATTCGACCAAAATCGAAGTGGAAAATGTTATGG 240
 DB |||||
 181 CCAACAGGC-----GACTTATATGACCTTACCATACTAGTTTAAAGAAATATTTT 231
 QY 241 GCGGGGATCTGAGGGGTTCTGTCAAAACCTCC-----TTATCTTAACAGCTGGGC 294
 DB |||||
 232 GGTGGCGATTGGCAGGGATTAATTAACAAATTAATGATGGTTATCTTACTGTTATGGGA 291
 QY 295 GTAACGACAAATCTGTTCTCCCGGTTTTGGCAATCTGGATACACT-----G 342
 DB |||||
 292 ATTACAGTATATGATTTGCAACCTGTAGAAACATTTACGCAGTTTTCGCAGATCC 351
 QY 343 GCGGGACCGATPAACACGGGTATCTACGGATCTGACGCGGATTTTAAACAGATGAG 402
 DB |||||
 352 ACTTTGGCGGAAGCACATCTATCATGTTTACTGGGCAGCAGACTTCAAAAAGACAAAT 411
 QY 403 GAACATTTCCGGAATTTGACACACATTTGACACGTTGGTCAATGATCTCACCACAAACGA 462
 DB |||||
 412 CCTTTTGGAGCTTTACAGATTTCAAAATCTCATAGCAAGCTCATGCTCAAT 471
 QY 463 ATCAGGTGATTTGCTGTTTGGCCCAATCATTCGACTCCCTTTTAAAGCAAAACGATTCC 522
 DB |||||
 472 ATAAAAGTTAATATAGATTTTGCACCAATCATATACATCTCTCGATCAGAGACAGACCT 531
 QY 523 ACCTTTGCGGAAGCGGCGCTCTACACATATGGAACCTATATGCGCAATTTATTTGAT 582
 DB |||||
 532 ACCTATGGGGAATTTGGTAGATTATATGACAAATGAGGATTTACTTGGTGGTTATACCAAT 591
 QY 583 GAGCCACAAAAGGTTACTTCCACCAATAATGGGACATCAGCAACTGGGACGACCGGTAC 642
 DB |||||
 592 GA---TACAAATGGATATTTTCATCATATATGGAGNACTA-----AT 630
 QY 643 GAGCGCAATGGAATAAATTCACGGATCCAGCGGTTTCTCGCTTGGCGAATTTTGGCAG 702
 DB |||||
 631 TTTTTCATCATATCAAGATGGAATTTACCGTAATTTATTTGACTTAGCAGATTTAGATCAG 690
 QY 703 GAAAATGGCAGGATGCTCAATACCTGACCGATGCGGGTTTCAATTTGTTAGCACATGA 762
 DB |||||
 691 CAGATAGCACTATTGATTCATATTTAAAGCGGCAATTAACACTAGTTAGTGGTATGGGG 750

QY 763 GCGGATGCTTTGCGGATTTGATGCGGTGAAGCATTTTAAATTCGGGGTTCTCCAAATCGTGTG 822
 DB |||||
 751 ATTGATGTTATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
 QY 823 GCGGATAAATCTGTACCAAAAGAAAGACATTTTCTGCTGGGGCAATGTTACGGAGATGAC 882
 DB |||||
 811 ATGGAATCTATTTTAAGTTATAGACAGTTTTCATTTGGCGAGTGGTACCTTGGAAACC 870
 QY 883 CCCGGAACAGCCCAATCATCTCGGAAAGGTCCGGTACCCCAACACAGCGGTCTCAATGTG 942
 DB |||||
 871 AATGAAGTAG-----ATCCTAATAATACGTATTTCCTCAATGAAAGTGTATGAGCCTT 924
 QY 943 CTGGATTTTGTATCTACACAGCGGTGTTGAAAATGTTTTCGGCACATTTACGCAAAACGATG 1002
 DB |||||
 925 CTGATTTTGTAGTTTGTCTCAAAAAGTTCGTCAAGTATTTACAGACAAATACAGACACTATG 984
 QY 1003 TAGCATCTTAACAATATGTTGAACCAAAACGGGAAACGAGTACAAATACAAAGAAAATCTA 1062
 DB |||||
 985 TATGACTTGTATTCGATGATTCAGTCTACTGACGAGATTTAATTTTCAATAATGATATG 1044
 QY 1063 ATCAATTTATCGATACCAATGATATGTAAGATTTCTTTTCGGTAAATTCGAACAGCGC 1122
 DB |||||
 1045 GTTACATTTATAGATAATCATGATGACAGATTTTATACAGGA---GSCAGTACACGG 1101
 QY 1123 AATTTGCACACAGCGCTTCTCTTCTTCTTCTTCTGCGGGGTACGCCCTCCATCTATTAT 1182
 DB |||||
 1102 CTGTTGAGCAAGCTTTAGCATTTTAACTTCTCGCGGTGCTCTCTATATATTTAT 1161
 QY 1183 GGAACCGAACAAATACATGCGCGCGGCAATGATCACCCGTACCAACCGGGGATGATGCGCGC 1242
 DB |||||
 1162 GGTACAGAGCAATATATGACAGTAAATGAGACCTTTATATATAGACTATGATGACGTCA 1221
 QY 1243 TTTGATAGCAACACCGCGCTTTTAAAGAGTGTCACTCTGCGGGGGTTCGCGAGGAAC 1302
 DB |||||
 1222 TTTGACACCAACAGCGGCATATAATGTGATAAAAAAGCTTCTCCACTGCGTAAATCT 1281
 QY 1303 AATGCGGATCTCAGTACGCGCACCCAGCGGTGGATCAACATGATGATGATTTTACAT 1362
 DB |||||
 1282 ACCCTGCAATTTGTTACGTTACAGTACAAACAGAGGATGATTAATATGATTTTACAT 1341
 QY 1363 TATGAACGGAATTTTCAACGATGCTGTTGGTGGCCATCAATCGAAACACGCAATCC 1422
 DB |||||
 1342 TATGAAGACCAATTTGTTAATACGTTGCTCTCTGTTGCTATTAATCGTAACTTTTCAAG 1401
 QY 1423 TCTATTTCGTTTTCGCGTTTCGAGCGGCTTGCCTGCAATGCGAGCTATCGGATTAATCTG 1482
 DB |||||
 1402 AGCTATTATCATTCGCGCTTGTACACCGCATTTGCTCGGGAAACATATTTCTGACATGCTT 1461
 QY 1483 TCAGGCTGTTTGGGGGAAACGGGATTTTCGTTTC---CAATGGAAGTGTGCTTTCGTTTC 1539
 DB |||||
 1462 GCGGATTTTAAATGCGAGTAGTATTACAGTACTAGTAAATGGTCTGTACACCGTTT 1591
 QY 1540 ACCTTTCGCTCGGAGCGGTTCTGTTTGGCAGTACAGCAATCCGCTTTCAGCGCGCA 1599
 DB |||||
 1522 ACCCTTTCGCTGCTGAGTTGCGAGTATGCGAGTATGCTGAGTAACTAACTAACTCCTCAATG 1581
 QY 1600 ATCGGATCGGTTCTCCAAATATGGGGATTCGGGTTATGTTGTTACGATCAGTACAGCGGAAA 1659
 DB |||||
 1582 ATAGACATGTAGACCGCAATGACAAAGGACGAGGAGCATTAACCATAGATGGAAG 1641
 QY 1660 GGTTTTGGGACGACGAGGGAACCGGTGACATTTTCGGGAGTGAACGAGCTGTGAAATCC 1719
 DB |||||
 1642 GATTTTGGCAACACAGCAGGTCAAGTATTTTGGGACAACTCTCTGCACTATTGTTGCA 1701
 QY 1720 TGAACATCCAAATGGAATGAGTACGTTTCCCAACATGCGCGCGGGCTGACCGATG 1779
 DB |||||
 1702 TGGGAAGATACCTGAAGTAAAGTAAAGTTCTCTGTTTAACTCTCTGAAAAATATAACATT 1761
 QY 1780 AAGT---CACCGGGGTGGAGTTTCCAGCAATCTGAT---TCTTACAAATTTTGGAT 1833
 DB |||||
 1762 ACATTAACAAACAGCATCAGGAGTTTCAAGCAATAGTATTAACAAATATCAATGTTTAAACG 1821

```

QY 1834 GGAACGAGACATCGGTTGCTTTACTGTAAGTGCCTCCGACCACTGGGGAT 1893
Db 1822 GGAATCAGGTATGTTAGATTTGTAGTAATAATG---CTACACCGTGTGGGAGAA 1878
QY 1894 AAGATTACCTGACGGGCAACATACCGGAATTGGGAATTGGAGCAGCATGACGCGA 1953
Db 1879 AATGATATCTTACGGGCAATGAGCTGAACCTTGGCACTGGGATACATCGAAGGCAATA 1938
QY 1954 GCGTTTAACTGCGCAAGGCGCCCTGCTGCGCCCAATATCCGGATTCGGTTTATGTA 2013
Db 1939 GGAACCAATGTTTAAACAGG-----TTGTGTATCAATATATCCATGCTATACGAT 1989
QY 2014 TTACGCGCTCCACGAGGAAGACGATTCAATCAAGTTCTTCAATCAAGCGTGGGATGA 2073
Db 1990 GTAAGTGTGCTGCTGTTACTATATAGTATTAAGTTTAAAGAAATGG---TAGT 2046
QY 2074 ACGATTCAATGGGAGAAATGTTGGAACCAAGTGGCCCACTCCACGGGTGCAACCGT 2133
Db 2047 ACTGTAACTGGGAAGGTGGATACCAACCAAGTATATACTACACCCACTTCTGTGATGCT 2106
QY 2134 AACATTACTGTTACGTGGCAA 2154
Db 2107 ACTGTATTTAGACTGGCAA 2127

```

RESULT 6

```

US-08-204-656B-9
; Sequence 9, Application US/08204656B
; Patent No. 5538882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuo
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; APPLICANT: Honda, Koichi
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
; TITLE OF INVENTION: Oligosaccharide Using The Enzyme
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,656B
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:

```

```

; LIBRARY: Library of chromosomal DNA from Bacillus
; LIBRARY: macerans, pMAC, generated by treating chromosomal DNA from Bacillus mac
; LIBRARY: IAM1243 with a restriction enzyme, and inserting and linking restriction
; LIBRARY: fragments to pBR322
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2061
; OTHER INFORMATION: /note= "Nucleotides 1-2061
; OTHER INFORMATION: correspond to nucleotides 82-2142 of the Bacillus macerans
; OTHER INFORMATION: cyclomaltodextrin glucanotransferase structural gene"
US-08-204-656B-9

Query Match 13.7%; Score 296; DB 1; Length 2061;
Best Local Similarity 51.5%; Pred. No. 4.7e-86;
Matches 1067; Conservative 0; Mismatches 950; Indels 73; Gaps 14;

QY 116 TCNAGGGGAGGAGTATTTACGAGATATCATGACCGGTTTACGATGGGACACGACGA 175
Db 35 TCAGTACGAGGAGTATTTACGAGATATGACCGGCTTCGCGACGGGACGACGACGA 94
QY 176 ACAACAATCTCTGCCAAAGTTATGGACTTTACGATCCGACCAAAATCGAAGTGGAAAATGT 235
Db 95 A---CAATCCGCGGGGAGTGGTTTCAGCGGCGACCGATCCAATTTGAAGCTCTATTCG 151
QY 236 ATTGGGGCGGGATCTCGAGGGGTTTGTCAAAAATCTTCTTATAAAGCTGGGCG 295
Db 152 GGGGAGACTGGCAGGGGATTTATCGAAGATTAAACGACGGTTATTTGACCGCATGGGCG 211
QY 296 TACGACAAATCTGTTGTCCTCCCGGTTTGGCAATCTGGATACACTGG-----CGG 346
Db 212 TCACCGCCTCTGGATATCCCACTGTGGAAATATCACTCCGTCACTAAGTATTCG 271
QY 347 GCACCGATACACGGGCTATCAGGATCTCGACGCGGATTTTAAACAGATGAGGAAC 406
Db 272 GCGTTAAACATACCTCTTATACGGTTGGTGGGAGGGATTTTAAAGCAAAACACGACG 331
QY 407 ATTTGGGAATTTGACACACATTTGACAGTTTGCATGATGCTCACCACAAACGGAATCA 466
Db 332 CTTTCGGGATTTTGGCGATTTTCAAAATCTGATGAT-ACGCTCAGCTCATACATCA 390
QY 467 AGGTGATTTGCGACTTTTGTGCCCAATCATTCGACTCCTTTTAAAG--GCAACGATTCAC 524
Db 391 AGGT-CGGATCGACTTCGCCCCCAACCAAGTCTCCGCGGAGGAGGACGACCCCGG 449
QY 525 CTTTGGGAGGCGGCGCCTCTACAGATGGAACCTATATCGGCAATTTTGTATGA 584
Db 450 CTTGCGCGAGAACGGTGGCTGTATGATACGGTTGCTGCTCGCGGCTACAGCAATGA 509
QY 585 CGCAACAAAAGGGTACTTCCACCATATATGGGACATCAGCAACTGGGACGACCGGTACGA 644
Db 510 TAGGGC---GGCCTTTTCCATCATACGGGGGACCGATTTT----- 550
QY 645 GGCGCAATGGAAGAACTTTCAGGATCCAGCGGTTTCTCGCTTGGCGATTTGTCAGGA 704
Db 551 --CCACGATTGAGACGGGTATTTACAAGAACCTCTACGACCTCGCGGACATCAACATAA 608
QY 705 AAATGGCACGATTGCTCAATACCTGACCGATGCGCGGTTCAATTTGATGACATGAGC 764
Db 609 CAACAGCTATGAGACGCTTATTTTAAAGCGCTATCGACCTTTGCTCGCATGGGTGT 668
QY 765 GATGTTTGGGATTTGATGCGGTGAAGCATTTTAAATTCGGGGTTCTCCAAATCGTTGG- 823
Db 669 GAGCGGATTCGTTTTCACGGGGTGAAGCAGTATCTTTGGGCTGSCAAAAGGCTTCGT 728
QY 824 --CCGATAACTGTACCAAAAGAAAGACATTTTCTGTTGGGGATGTTAGCGGATGA 881
Db 729 TTCTCGATTTACGGGCGGATCATCCGGTATTTACGTTTCGGGGAATGTTATCTTTGGCG 788
QY 882 CCGCGGACAGCAATCATCTGGAAAGTCCGGTACGCCAACACAGCGGTGTCAATGT 941
Db 789 GATCAACCG-----ACGGAGACACATTAATTTGCCCAACGAAGCGGATGACCT 842
QY 942 GCTGGATTTTGTATCTCAACACCGGTGATTCGAATGTTTCGGCACATTTTACGAAACGAT 1001

```

Db 843 GCTGACCTTTGAAATACGCGCAGGAAGTGGCGAAGTGTTCGGGACAAACCGGAAACGAT 902
Qy 1002 GTACGATCTTAAACATATGCTGAACCAACCGGGAAGAGTACAAATACAAAGAAATCT 1061
Db 903 GAAGGATCTCTATGAGGTGTGCGCCAGCAGGAGTTCGCAATACGACTACATCAATAT 962
Qy 1062 AATCAGATTTATCGATACATGATGTCAGATTTCTTTTCGGTAAATTCGACAGGC 1121
Db 963 GGTGACCTTTCATCGACACCATGATATGACCGGTTCAGGTTGCCGGTTCGGTACGCG 1022
Qy 1122 GAATTTGACACGAGCGCTTGCTTCTCACTTCGCGGGTACGCTCCATCTATTA 1181
Db 1023 GGCACGAGCAGCGGTGGGTGCTGAGCTGACCTTCGCGGGTGCAGCCTACTTA 1082
Qy 1182 TGAACCGAACATATATGAGCGGCAATGACCCGTPACAAACCGGGGATGATCCGGC 1241
Db 1083 CGGACGAGCAGTACATGACCGCGGATGGCAGCCCAACAAACCGGGGATGATGACCTC 1142
Qy 1242 GTTTGATAGCACACCCCTTTAAAGAGGTGTCACTCTGCGGGGTTCGCGAGGA 1301
Db 1143 GTTTAATACCGGACGAGCGCTTATAAGTGATTCAGGCAATGGCGCGCTCGTAAATC 1202
Qy 1302 CAATCGCGGATTCAGTACGCGCACCAACCGAGGTGATCAACAAATGATGTTTACAT 1361
Db 1203 CAATCGCGCATCGCTTATGGAGCAGCAGCAGCGCTGGGTAAACAGATGTGTGAT 1262
Qy 1362 TTATGAAGGAAATTTTCAACATGTGTTGTGGTGGCCATCATCGAAACACCAATC 1421
Db 1263 TATTGAACCAATTTGCGCAGCAGCGCGCTTTGTTGCGGATTAATCGAAACTCTCGCG 1322
Qy 1422 CTCCTATTCGATTCGCTTTGAGCAGCGCTTGCAAAATGGCAGCTATGCGGATATCT 1481
Db 1323 CGTATTCGATTTCCGGTCTGTGTGATTCGCTGGCGGGGACCTTATTCGGATGAT 1382
Qy 1482 GTGAGGCTGTGTTGGGGGGAACGGGATTTCCGTTTCCAAATGGAAGT---GTGCTTCGTT 1538
Db 1383 GAACGGAATCTTAAACGCAACTCCATTACCGTGGCAGCGCGCGCGCTCACCACCTT 1442
Qy 1539 CAGCTTTGGCTGGAGCGCTGCTGTTGGCAGTACGACATCCGCTTCAGCGCGCA 1598
Db 1443 TAGCGTGGCGCGCGCGCGGATGCGGATGCGGATGCGGCGGAAACGCTCGCGCG 1502
Qy 1599 AATCGGATCGTGTCTCAATATGCGGATTCGCGGTAATGTGTCAGATCGAGGAA 1658
Db 1503 GATCGCATGTGCTGCTCCACCATGGCGCAGCGGGGATATAGTACGATTCAGCGCG 1562
Qy 1659 AGTGTGAGCAGCGAGGACCGTGACATTTGCGGAGTGACGCGACTG----- 1711
Db 1563 CGCTTTGCGCGCAGCGCGGCGCGGTTTATTTGCGGACGACGCGGTGACCGCTCGCG 1622
Qy 1712 --TGAATCTCGACATCCATCGATGAGTGTACGTTCCCAACATGCGCGCGGCT 1769
Db 1623 CATCGTAACTGGAGGACGACGATTAAGCGGTCTACCGAGGTGCGCGGGGAA 1682
Qy 1770 GACCGATGGAAGTCT---ACCGCGGTGAGTTTCCAGAAATCTGATTTTCAATAT 1826
Db 1683 AACCGGCTATCGGTCAAAACGTCGTCGCGCACCGCAGCAATACATTTAAAGTTCAA 1742
Qy 1827 TTTGAGTGGAAACGACAGCATCGGTTGTTTACTGTGAAAAGTGGCTTCGACCACT 1886
Db 1743 TGTACTGACGGGGATCAGGTACGCTGGCTTTCTTGTGTCATCAAGCAATACCAATTA 1802
Qy 1887 GGGGGATTAAGTTTACCTGAGCGGCAACATACCGGAATTTGGGGAATTTGGAGCAGGATAC 1946
Db 1803 CGGAACAAATGTTTATCTTGTGCGCAGCGCGCGAGCTCGGACCTGGGACCGGAA 1862
Qy 1947 GAGCGGAGCGTTTAAATGCGGAGGCGCGCTGCTGCGGCGCCAAATATTCGATTTGTT 2006
Db 1863 AGCGATTTGGCGCGGATGTAATCAATCAGGTGATCG-----CCAAGTACCGCTCTGTA 1913
Qy 2007 TTATGTTATTCAGGTTTCAGCAGGAAAGAGATTCATTAAGTTTCTTCATCAAGCGTGC 2066

Db 1914 TTACGATGTCAGGTGCGCGGGGACAAAGCTGGATTTTAAATTTATTAAG--CG 1970
Qy 2057 GGATGGAACGATTAATCGGGAATGTTTCGAACCACTGGCCCAACTCCCAAGGTGC 2126
Db 1971 CGCGGTGACGTGACTTGGGAAGCGGGGCAACCATACGTACACGCGCGGCGGCG 2030
Qy 2127 AACCGGTACATTAATGTTTACGTTGGGAAA 2156
Db 2031 CGTAGGACGTTACGTTGGGACTGGCAAAA 2060

RESULT 7
US-08-470-702-5
; Sequence 5, Application US/08470702
; Patent No. 5831149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
US-08-470-702-5

Query Match 13.7%; Score 296; DB 1; Length 2061;
Best Local Similarity 51.5%; Pred. No. 4.7e-86;
Matches 1067; Conservative 0; Mismatches 930; Indels 73; Gaps 14;

Qy 116 TCAGAGGGGAGCTGATTTACGAGATTATCATTCGCGTTTACGATGGGACACGACGA 175
Db 35 TCAGTACGGAGCTCATCTATCATGATTGTGACCGGCTTCGCGACGGGACGACGA 94
Qy 176 ACAACATCTCTGCAAAAGTTATGACTTTACGATTCGACCAAAATCGAAGTGGAAATGT 235
Db 95 A---CAATCCGGGGGATGCGTTTCAGCGGCGGACCGATCCATTTGAAGCTCTATTTCG 151

QY 236 ATTGGGCGGGGATCTGGAGGGGTTCTCAAAATCTCTTATCTTAAACAGCTGGGCG 295
Db |||||
QY 152 GGGGAGACTGGCAGGGGATATCGACAAGATTAACGAGGTATTATTGACCGGATGGGCG 211
Db |||||
QY 296 TAAACGACAATCTGGTTGTCCCGGTTTGGACATCTGGATACACTGG-----CGG 346
Db |||||
QY 212 TCACCGCCCTCTGGATATCCCAACCTGTGGAAATATACCTCCGTCATCAAGTATTCG 271
Db |||||
QY 347 GCACCGATACACGGGCTATACGATACGATCTGGACGCGGATTTTAAACAGATTGAGAAC 406
Db |||||
QY 272 GCGTTAAACAATGCTCTTATCACTGTTGCTGGCGAGGGATTTTAAAGCAAAACCAACGAG 331
Db |||||
QY 407 ATTTTCGGGAATTTGGACCAATTTGACGCTTTGGTCAATGATGCTACCAAAACGGAATCA 466
Db |||||
QY 332 CTTTCGGGATTTGGCGATTTTCAAAATCTGATTTGAT-ACGCTCACGCTCATTAACATCA 390
Db |||||
QY 467 AGGTGATGTGCACTTTGTGCCCAATCATCTGATCTCTTTTAAAG--GCAACGATTTCCAC 524
Db |||||
QY 391 AGGT-CGGATCGACTTGGCCCAACACACGCTCTCGGCGGAGCGGAGCAACCCCGG 449
Db |||||
QY 525 CTTTTCGGGAAGGGCGGCTCTCAACAATGGAACCTATATGGGCAATTTTGTATGA 584
Db |||||
QY 450 CTTTCGGGAGAACGGTGGCTGTATGATGATGATGATGATGATGATGATGATGATGATG 509
Db |||||
QY 585 CGCAACAAAGGGTACTTCCACCAATATGGGGACATCAGCAACTGGGACGACCGGTACGA 644
Db |||||
QY 510 TACGGCC---GGCCTTTTCCATCATAAACGGGGGACCGATTTT-----550
Db |||||
QY 645 GGGCAATGGAATACTTACGATCCAGCCGCTTTCTCGCTTCCGATTTGTCGAGGA 704
Db |||||
QY 551 --CCACGATTAAGACGGTATTTTACAAGAACCTCTACGACTGGCGGACATCAACCATAA 608
Db |||||
QY 705 AAATGGCAAGATGCTCAATACCTGACGATGCGCGGCTTCAATTTGTGTAGCACATGGAGC 764
Db |||||
QY 609 CAACAACCTATGACGCTTATTTTAAAGCGCTATCGACTTTGCTCGGATGGGTGT 668
Db |||||
QY 765 GATGGTTTGGGATTTGATCGGCTGAGCATTTTAAATTCGGGCTTCTCCAAATCGTTGG- 823
Db |||||
QY 669 GGACGGGATTCGTTTTCAGCGGGTGAACGATGATCTTTTCGGCTGCAAAAAGCTTCGT 728
Db |||||
QY 824 --CCGATAACTGTACCAAAAGAAAGACATTTTCTGCTGGGGAATGTACGAGATGA 881
Db |||||
QY 729 TTCCTCGATTTACGGCGGATCATCCGGTATTACGTTTCGGGAATGTATCTTTGGCGC 788
Db |||||
QY 882 CCGCGGAACGCCAATCATCTGAAAGGTTCGGTACGCCAACAACAGCGGTTCATGTT 941
Db |||||
QY 789 GGATCAAAACCG-----ACGGAGACCAACATTAATTCGCCAAGCAAGCGGATGAACCT 842
Db |||||
QY 942 GCTGGATTTGATCTCAACACGGTGTATCGAAATGTCTTCGGCACATTTACGCCAACGAT 1001
Db |||||
QY 843 GCTGGACATTTAAATACCGCAGGAAGTTCGCGAAGTGTTCGGGACAAACCGGAACGAT 902
Db |||||
QY 1002 GTACGATCTTAACATATGTGAAACCAACCGGGAACGAGTACAAATACAAAGAAATCT 1061
Db |||||
QY 903 GAAGGATCTCTATGAGTGTGCGCAGACGAGTTCGCAATACGACTACATCAACAATAT 962
Db |||||
QY 1062 AATCACAATTTGATTAACATGATATGTCAAGATTTCTTCGGTAAATTCGAACAGGC 1121
Db |||||
QY 963 GGTGACCTTCATCGACACCAATGATATGGACCGGTTCCAGGTTGCGCGTTCGGGTACGCG 1022
Db |||||
QY 1122 GAATTTGACACGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1181
Db |||||
QY 1023 GGGCAGCGGAGAGCGTTGGCGCTGACGCTGACTTTCGCGCGGCTGGCAGCATCTACTA 1082
Db |||||
QY 1182 TGAACCGGAACATATACGACGCGGCAATGACCCGTPACACCGGGGATGATCCCGGC 1241
Db |||||
QY 1083 CGGACGAGGAGTACATGACCGGCGATGGCAGCCCAACAACCGGCGCATGATGACCTC 1142
Db |||||
QY 1242 GTTTGATACGACCAACCGGCTTTTAAAGAGGTGTCACTCTGCGGGGTTGCGGAGAA 1301
Db |||||
QY 1143 GTTTAATACGGGACGAGCGCTTATAAGTGTTCAGGCATTTGGCGCGCTCGCTAAATC 1202
Db |||||

QY 1302 CAATGGCGGATCCAGTACGACACCCAGCGGTTGGATCAACCAATGATGTTTACAT 1361
Db |||||
QY 1203 CAATCGGCGCATCGCTTATGGAGACGACAGAGCGCTGGGTTAAACAAGATGTTGAT 1262
Db |||||
QY 1362 TTATGAACGGAATTTTCAACGATGTCGTGTGTGGCCATCATTCGAAACACGCAATC 1421
Db |||||
QY 1263 TATTTGAACCAATTTGGGACGACAGCGCGCTTTGGTGGCGATTAATCGAAACTCGCTCCGC 1322
Db |||||
QY 1422 CTCTATTCGATTTCCGCTTTGCGACGCGCTTGCCTTGGCAGCTATGCGGATTTCT 1481
Db |||||
QY 1323 CGCTTATCCGATTTCCGCTTGTGTGCTGCTGCGCGGCACTTTATTCGATGTTAT 1382
Db |||||
QY 1482 GTACGCGCTTTTGGGGGGAACGGAATTTCCGTTTCCAAATGGAAGT---GTGCTTTCGTT 1538
Db |||||
QY 1383 GAACGGACTCTTTAAACGGCAACTCCATACGCTGGGCGAGCGCGCGCGCTCACCAACTT 1442
Db |||||
QY 1539 CAGCTTTCGCTGAGCGCTGCTGTTTGGCAGTACGACACATCCGCTTCAGCGCGCGCA 1598
Db |||||
QY 1443 TAGCTGGCGCGCGCGCGGACGCGGCTATGGCAGTACACGCGCGGAAACGCTCGCGCGC 1502
Db |||||
QY 1599 AATCGGATCGGTTGCTCCAAATATGGGGATTTCCGGGTAAATGTGTACGATCGACGGAA 1658
Db |||||
QY 1503 GATCGCAATGTGGTCCCACTATGGCGCAGCGCGGGAATATATGACGATTCGACGCGC 1562
Db |||||
QY 1659 AGTTTGGGACGACGAGGAACCGTGTACATTTGCGGGAGTGTACGAGCTG-----1711
Db |||||
QY 1563 CGGCTTGGCGGACGCGCGGCAAGGTTTATTTTCGGGACGACGCGCGTACCGGCTCGG 1622
Db |||||
QY 1712 --TGAAATCTCGACATCCAAATCGGATTAAGTTGATCTTCCCAATGCGCCCGCGGCT 1769
Db |||||
QY 1623 CATCGTAACTGGGAGGACGACGAGATTAAGCGGCTCATACGAAAGTTCGCGCGGCGCA 1682
Db |||||
QY 1770 GACCGATGTGAAAGTC--ACCGCGGTGGAGTTTCCAGAACTGTATTCTTACAAATAT 1826
Db |||||
QY 1683 AACGGCGTATCGGTCAAAACGTCGTCGCGCACCGCGCAATATACATTTCAAAAGCTTCAA 1742
Db |||||
QY 1827 TTTGAGTGGAAACGACAGACATCGGTTGTGTTTACTGTGAAAGTGGCTTCGACCAACCT 1886
Db |||||
QY 1743 TGTACTACGCGGGGATTCAGGTCACGCTGCGGTTCTTCTGTCAATCAAGCAATACCAATTA 1802
Db |||||
QY 1887 GGGGGAATGAATTTACCTACGCGGCAACATACCGAAATTTGGGGAATTTGGAGACGGATAC 1946
Db |||||
QY 1803 CGGAACAAATGTTTATTTCTTGTTCGCAACGCGCGGAGCTTCGCACTGGGACCGGAACAA 1862
Db |||||
QY 1947 GAGCGGAGCTTAAACATGCGGAAGGCGGCTGCTGCGCGGCAATATTCGATTTGTT 2006
Db |||||
QY 1863 AGCGATTTGGGCGGATGATCAATCAGGTATCG-----CCAGTACCGCTCTCGTGA 1913
Db |||||
QY 2007 TTATGATTTCCAGGTTTCCAGCAGGAAGACGATTCATTTCAAGTTCTTTCATCAAGCGTGC 2066
Db |||||
QY 1914 TTACGATGTCAGCGTCCGCGGGGCAAAAGCTGGATTTTAAATTTTAAAG---GG 1970
Db |||||
QY 2067 GATTTGAACGATTTCAATCGGGAATGTTTCGAACCACTGGGCAAACTCCCAAGGCTGC 2126
Db |||||
QY 1971 CGCGGCTACGCTGACTTGGGAAGCGGGGCAACCATACATACGACGCGCGCGGCGGCG 2030
Db |||||
QY 2127 AACCGTTAACTTCTGTTTACGTCGCAAA 2156
Db |||||
QY 2031 CGTAGGACGCTGACGGTGGACTGGCAAAA 2060
Db |||||

RESULT 8

US-08-467-831-5

; Sequence 5, Application US/08467831

; Patent No. 5635378

; GENERAL INFORMATION:

; APPLICANT: NATSUI, IKUO

; APPLICANT: ISHIKAWA, KAZUHIKO

; APPLICANT: MIYAIRI, SACHIO

; APPLICANT: KONDA, KOICHI

; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,

; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING

; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME

; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 ; STREET: 8110 Gatehouse Road, Suite 500 East
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22042
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/467,831
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/204,656
 ; FILING DATE: 02-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WEINER, MARC S.
 ; REGISTRATION NUMBER: 32,181
 ; REFERENCE/DOCKET NUMBER: 234-252P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 205-8000
 ; TELEFAX: (703) 205-8050
 ; TELEX: 248345
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2061 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (synthetic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; US-08-467-831-5

Query Match 13.7%; Score 296; DB 1; Length 2061;
 Best Local Similarity 51.5%; Pred. No. 4.7e-86;
 Matches 1067; Conservative 0; Mismatches 930; Indels 73; Gaps 14;

Qy	116	TCAAAGGGGCGCTGATTACAGATTATCATTTAGCCGGTTTACGATGGGGACAGACGA	175
Db	35	TAGTACGGAGTCACTTATCAGATTGTGACCGACCGCTTCGGCGGAGGGGACAGCA	94
Qy	176	ACAAACAATCTGCGCAAAAGTTATGGACTTTACGATCCGACCAATCGAATGGGAAAATGT	235
Db	95	A---CAATCCGGCGGGGGATGCGTTTCAGCGCGACCGATCCAAATTTGAAGCTCTATTTCG	151
Qy	236	ATTGGGCGGGGATCTCGAGGGGTTCTGTCAAAACITCTTATCTTAAACAGCTGGCG	295
Db	152	GGGGAGACTGCGAGGGGATTATCGACAGATTACGACGGTTATTGACCGGCAATGGCG	211
Qy	296	TAAACACAATCTGGTTGTCCCGCTTTTGGACAATCTGGATACACTGG-----CGG	346
Db	212	TCACCGCCCTCTGGATATCCCACTGTGGAAATATCACTCCGTCACTAATTCG	271
Qy	347	GCACCGATAACACGGGCTATCAGGACTGGAACCGCGGATTTTAAACAGATTGAGAAC	406
Db	272	GGGTTAAACAATACACTCTTATCAGCGTTGGTGGGGGAGGGATTTTAAACAAACCAACGACG	331
Qy	407	ATTTCGGGAATTGACCAACATTTGACAGCTTGGTCAATGATGCTCAACCAACGGAATCA	466
Db	332	CTTTTGGGGAATTTGCGGATTTTCAAAATCTGANTGAT-ACGCTCAGCTCATACATCA	390
Qy	467	AGTGATTTGCGACTTTTGTCCCAATCAATTCAGCTCCCTTTTAAAG--GCAAAAGATTCCAC	524
Db	391	AGGT-CGGATCGACTTCGCCCCCAACCAACAGCTCTCCGGCGAGGAGGACGAAACCCGG	449
Qy	525	CTTTTCGGAAGGCGGCGCCCTCTACAACAATGGAACTTATATGCGCAATTTATTGTA	584

Db	450	CTTCCCGCAGAACGGTGGCGCTGTATGATAAAGGTTTCGCTGCTCGCGCGCTACACCAATGA	509
Qy	585	CGCAACAAAAGGTACTTCCACCATAATGGGGACATCAGCAATCGGACGACCGGTACGA	644
Db	510	TACGGCC--GGCCTTTTCCATCATACGGGGGACCGATTTT-----	550
Qy	645	GGCGAATGGAAAACCTTCAGGATCCAGCCGGTTCTCGCTTCCGATTTGTGCGAGGA	704
Db	551	--CCAGATTGAAGACGGTATTTTCAAGAACCCTCTACGACCTGGGGGACATCAACCATTA	608
Qy	705	AAATGGCAGATTGCTCAATACCTGACCGATGCGGGTTCAATTGGTAGCACATGAGC	764
Db	609	CAACAACGCTATGACGCTTATTTTAAAGCGTATCGACCTTTGCTCGCATGGGTGT	668
Qy	765	GGATGGTTTCGGATTGATGGGTGAAGCAATTTAATTCGGGTTCTCCAAATCGTTGG-	823
Db	669	GGACGGATTTCGTTTTCAGCGGGTGAAGCAGTATCTTTCCGGCTGGCAAAAAGCTTCGT	728
Qy	824	--CCGATAAACTGTACCAAAAAGAAAGACATTTTCTGTTGGGGGAATGTACGGAGATGA	881
Db	729	TTCTCGATTTCAGCGGGGGATCTCCGGTATTTCGTTTCGGGGGAATGGTATCTTGGCGC	788
Qy	882	CCCCGAAACAGCCAAATCATCTGGAAAAGGTCCGGTAGCCCAACAAACAGCGGTGCAATGT	941
Db	789	GGATCAAAACG-----ACGGAGACAACATTAAATTCGCCAACGAAAGCGGATGAACCT	842
Qy	942	GCTGGATTTTGATCTACACACGGTGATTCGAAATGTTTCGGGACATTTTCGCAAAAGAT	1001
Db	843	GCTGGACTTTGAATACGGCGAGGAAGTGGCGGAAGTGTTCGGGGCAAAAACGGAACGAT	902
Qy	1002	GTACGATCTTAAACAATATGGTGAACAAAACGGGGAACGAGTACAAATCAAAAGAAATCT	1061
Db	903	GAAGATCTCTATGAGGTGCTGGCCAGCAGGAGTCGAAATACGACTACATCAACAATAT	962
Qy	1062	AATCACATTTATCGATAACCATGATATGTCAAGATTCTTTTCGGTAAATTCGAACAAGC	1121
Db	963	GGTGACCTTCATCGACAACCATGATATGAGACCGGTTCCAGGTTGCGGCTTCGGTACGG	1022
Qy	1122	GAATTTGCACCGGCGCTGCTTTTCATCTCACTTCGCGGGGTACGCCCTCCATCTATTA	1181
Db	1023	GGGACCGCAGCAAGCGTTGGCGCTGACGCTGACTTCCCGCGCGTGGCCAGCATCTACTA	1082
Qy	1182	TGGAACCGCAACATACATGCGCAGCGCGCAATGACCGTACAAACGGGGGATGATGCCGC	1241
Db	1093	CGGACGAGCAGTACATGACCGCGATGGGACCCCAACACACCGGGCGATGATGACTC	1142
Qy	1242	GTTTGATACACAACAACACCGCTTTAAAGAGGTGTCAACTCTGGCGGGTTGGCAGGAA	1301
Db	1143	GTTTAATACCGGACGACGCGTTTAAAGTGATTCAGGCATTTGGCGCGCTGCGTAAATC	1202
Qy	1302	CAATGCGCGCATCCAGTACGACCCACCCAGCGTTGGATCAACATGATGTTTACAT	1361
Db	1203	CAATCCGGCGCTCGCTTATGGGACGACGAGCGCTGGGTTAAACAACGATGTTGAT	1262
Qy	1362	TTATGAACGGAATTTTCAACGATGCTGTGTGGTGGCCATCAATCGAAACACGCAATC	1421
Db	1263	TATTGAAGCAATTCGGCAGCAGCGCGCTTTGGTGGCGATTATCGAACTCGTCCG	1322
Qy	1422	CTCCTATTCGATTCGGTTTCAGACGCGCTTCGCAAAATGGCAGCTATGCGGATATCT	1481
Db	1323	CGCTATTCGATTCGGTCTGTGAGTTCGCTGCGGGGCGCACTTATTCGATGTTAT	1382
Qy	1482	GTACGGGCTGTGGGGGGAACGGGATTTCCGTTTCCATGGAAGT--GTCCGCTTCGTT	1538
Db	1383	GAACGGAATCTTAAACGGCACTCCATTTACGTGGGACGCGCGCGCGTACCAACTT	1442
Qy	1539	CACGCTTGGCGCTGAGCGCGTGTCTGTTTGGCAGTACAGCATCCGCTTCAGCGCGCA	1598
Db	1443	TAGCTGGCGCGCGCGCAGCGGGTATGGCAGTACACAGCGCGGAAACGTCGCGCG	1502
Qy	1599	AATCGATCGGTTGCTCCTAAATATGGGGATTCGGGTAAATGTTGTCAGATTCGACGGAA	1658
Db	1503	GATCGCAATGTGGTCCCACCATGGGCGGCAATATAGTACGATTCGACGCGCG	1562

```

QY 1659 AGGTTTGGGACGACGCGGGAACCGTGACATTTTGGCGGAGTGCACAGCGACTG----- 1711
Db 1563 CGGCTTTGGCGGACGCGCGGCGACCGGTTTATTTTCGGGACGACGCGGTTGACCGGTCGG 1622
QY 1712 --TGAATCTTGGACATCCCAATCGAATGGAATGAGTTTCCCAACATGGCGCGCGGCT 1769
Db 1623 CATCTAGCTGGGAGACACGAGATTAGGCGGTATACCGAAGTTCGGCGGCGCAA 1682
QY 1770 GACCGATGAAAGTC---ACCGCGGTGGAGTTTCAGCAATCTGTATCTTCAATAT 1826
Db 1683 AACCGCGGTATCGGTCAAAACGTCGTCGGGACCGCGCAGCAATACATTCAAAGCTTCAA 1742
QY 1827 TTTGAGTGAACGAGACATCGGTCTGTCTTACTGTGAAAGTGCCTCCGACCAACCT 1886
Db 1743 TGTACTGAGGGGATCAGTCTACGTGCTGTTCTGTCTCAATCAAGCAATACCAATTA 1802
QY 1887 GGGGATAGATTTTACCTGACGGGCAACATACCGGAATGGGAAATGGAGACGGATAC 1946
Db 1803 CGGACAAATGTTTATCTTCTGCGCAACCGCGGAGCTCGGACCTGGGACCGCGCAAA 1862
QY 1947 GAGCGGACCGTTAACAATGCGCAAGGCCCTGCTCGGCCCAATATTCGGAATGTTT 2006
Db 1863 AGCGATTGGCGGATGTACAAATCAGTGATCG-----CCAACTACCGCTCTGTA 1913
QY 2007 TTATGTATTCAGCGTTCCAGCAGGAAAGACGATTCAATTCAGTTCTTTCATCAAGCGTGC 2066
Db 1914 TTACGATGTCAGCGTTCGCGCGGGGACAAAGCTGGATTTAAATTTTATTAAG--GG 1970
QY 2067 GGTGGAAGGATTCAATGGGAGATGTTTCGAAACCGTGGCCACACTCCACGCGTGC 2126
Db 1971 CGCGGTACGCGTTCGGAAGCGGGGGGACCACTACGTAACAGCCCGCGCAGCGG 2030
QY 2127 AACCGTAACATTACTGTTACGTGGCAAA 2156
Db 2031 CGTAGGCGGTGACGCTGGACTGGCAAA 2060

RESULT 9
US-08-816-105A-2
; Sequence 2, Application US/08816105A
; Patent No. 5989882
; GENERAL INFORMATION:
; APPLICANT: Cruiger, Anneliese; Dellweg, Hans-Georg; Lenz,
; APPLICANT: Jurgens, Schroeder, Werner; Pape, Hermann;
; APPLICANT: Goeke, Klaus; Schaper, Beate; Hemker, Michael;
; APPLICANT: Piepersberg, Wolfgang; Distler, Jurgen;
; APPLICANT: Stratmann, Ansgar
; TITLE OF INVENTION: PROCESSES FOR PREPARING ACARVIOSYL
; TITLE OF INVENTION: TRANSFERASE AND FOR USING IT IN THE
; TITLE OF INVENTION: CONVERSION OF ACARBOSE HOMOLOGUES
; TITLE OF INVENTION: INTO ACARBOSE, FOR THE PREPARATION
; TITLE OF INVENTION: OF ACARBOSE HOMOLOGUES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sprung Kramer Schaefer & Briscoe
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: Apple Macintosh 6500
; OPERATING SYSTEM: System 7.5
; SOFTWARE: WordPerfect 3.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,105A
; FILING DATE: 14-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19625269.5
; FILING DATE: 25-JUN-1996

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19611252.4
; FILING DATE: 22-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: Bayer 9814-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-816-105A-2

Query Match 10.7%; Score 231.4; DB 2; Length 2582;
Best Local Similarity 49.8%; Pred. No. 7.1e-65;
Matches 1058; Conservative 0; Mismatches 951; Indels 116; Gaps 14;

QY 95 AAGCCAGCGATTCCGCAAGCGTCAAGGGGAGTGATTTACCAGATTATCATTTGACCGGT 154
Db 226 ACGCGGTAAACCGGCCAGTCTCAGCAGGACCTCTGCTACAGATCGCCACCGCGGT 285
QY 155 TTTCAGATGGGACACGACGAAACAATCTTCCCAAAAGTTATGGACTTTACGATCGA 214
Db 286 TCAGCAGCGGACCCCGCGGAAACAACAATCCGGGCAAGCTGCCCGCATTTCCCGGACA 345
QY 215 CCAATTCGA-----AGTGGAAATGTTTGGGCGGGGATCTCGAGGGG 259
Db 346 AGACCAAGCTGAACGACCGGAGGATGGCTCAATATCATGGAGGTGATTCGCGGCA 405
QY 260 TTCGTCAAAAATTCCTTATCTTAAACAGCTGGGCGTAAACCAATCTGTTGCCCGG 319
Db 406 TCACCCAGCGGATGGAGTACCTCAAGAACCTGGCGTTCGGCGCATCTGGATCTCGCCG 465
QY 320 TTTTGGACAATCTGGATACACTGGCGGACCGATTAACCGGCTATCAAGGATCTGGA 379
Db 466 ACGTCGACAACTCAACGTTCCGCGGAAACCGGCG---CACCGTTTACCGGCTACTGGC 522
QY 380 CGCGCGATTTTAAACAGATTGAGGAACATTTGGGAAATTTGGACCAACATTTGACACGTTGG 439
Db 523 CGCGCGACTTCAAGCGGCTCGAGAGCACTTCGACCGACGACGAGGTTTGAAGCGCTGG 582
QY 440 TCAATGATGTCAACAAACGGAATCAAGGTGATTTGATGTTTGGTCCCAATCATTCGA 499
Db 583 TGTCCGGCGGCGACGCCAGCAACATCAAGTGATCATGTGACTGGACGCCGACGCGACCA 642
QY 500 CTCCTTTTAAAGGCAACGATTCACCTTTTCGGAAGGGCGGCGCTCTACAAACAATGGAA 559
Db 643 ACCCGCGAA-----CCAGGCCGAGGACGCGCGCTCTACGAGATGGGC 687
QY 560 CCTATATGGCAATTTATTTGATGACGCAACAAAGGTTACTTCCACCAATATGGGGACA 619
Db 688 AGCTGTCGCGAGGTACGCGGCGGACAGTGCC---GGGCACTTCCACACCGCGCGCGCA 744
QY 620 TCAGCACTGGGACGACCGGTACGAGGGGCAATGGAAAACCTTACGGATCCAGCGGTTT 679
Db 745 TGGCGGACTTCAACGATCGCTTACCAGGACCACTACTACAGCTGGCGCA----- 793
QY 680 TCTCGCTTGGCGATTTGTTCGAGGAAATGGCAGATTGCTCAATACCTGACCGATCGG 739
Db 794 ----CATCGCGGACTCGACGACGAGAACCCGCGGTCGACCACTGCTCAAGACGAG 849
QY 740 CGGTTCAATTTGTAGCACAATGGAGCGGATGTTTGGGATTTGCGGTGAAGCATTTTA 799
Db 850 CCAACTACTGATGGACCGCGGGGTTCGACGCGCATCCGGGTTCGACCGCTCAAGCATGC 909
QY 800 ATTCCGGGTTCTCCAAATCGTTGGCCGCAATAATGTACCAAAAGAAAGACATTTTCTGG 859
Db 910 CGCTAGCTGGCAGCGGTCCTTTGCGCGCAACGCGGTCACTTCGCAAGAGCGCGGCATCT 969

```

QY 860 TGGGGAAATGGTACGGAGATGACCCCGGAAACAGCAATCATCTGGAAGAGGTTCGGTAAG 919
Db 970 TCGGCGAGTGTATCATGGGACCACTCGGATCCGCTTACGCCGACAGGTCAAGTTGG 1029
QY 920 CAAACAAAGAGGGTGTCAATGTGCTGGATTTTGAATCTCAACACGGTGATTCGAATGTGT 979
Db 1030 CCAACACAGAGGGATCGGGCCATGGAACTTCTACCAACCGCTCGATTCGGGACACCT 1089
QY 980 TGGGACATTTACGCAACAGATGTACGATCTTTAACTATATGTGAACCAACAGGGGAAG 1039
Db 1090 TCGCGGCG- - -CGGCTCGATGAAGTCCCTGGACGGCGATCACCAAGACCAACCGG 1146
QY 1040 AGTACAAATACAAAGAAATCTAATCAATTTATGATAACCAATGATATGTCGAATTC 1099
Db 1147 ACTACTCTACGAGCAGATCTGATCACTTCTTGGAACCAACGAGACACCGGGCTTGG 1206
QY 1100 TTTTCGTAATTCGAAACAGGCGAATTTGACCAAGGCGCTTGTTCATTCTCACTTCG 1159
Db 1207 GAGCGCTCAACAGGATTCGGCGCGCTGACCGGGCGCTGCGCTTCTGCTCACACCC 1266
QY 1160 GGGGTACGCTCTCATATATGAAACCGAAACAAATACATGCGAGC- - - - - 1206
Db 1267 GGGGTACGCTGCTGCTTCTACGGCAACCGAGCAGTACCTGCAACAGCACACCGTGAGG 1326
QY 1207 - - - - -GGCAATGACCGCTACAAACGGGGGATGATCCGGGTTTGATACGAA 1255
Db 1327 GCAGCAACAAAGGGAAGACCGGTACACCGGCCCGGATGGCGAGTTTCGACACCGACA 1386
QY 1256 CCACCGCTTTAAAGAGT- - -GTCAACTCTGGCGGGGTGGCGAGGAACAATCGGCCATC 1314
Db 1387 CGTTCGCTTACGGGAGATCCGGCGCTCTCCGACCTGCGCGGTGCAACCGCGGTG 1446
QY 1315 CAGTACGGCACACACCGAGGTGGATCAACAAAGATGTTTACATTTATGAACGAA 1374
Db 1447 GCTACGGGACCAACA- - -GCAGCGGTGATCAACAGCAGCAGTGTACTTACGAGCGCG 1505
QY 1375 TTTTCAAGATGCTGTTGGTGGCCATCAATCGAAACGCAATCCTCTATTTCGATT 1434
Db 1506 TTGGCGCAACAGCTGCTGACCGGCATCAACAGGCTCGCA- - -CGAGTACCGGCTC 1562
QY 1435 TCGGTTTCAGACGCGCTTCGCAATATGGCAGCTATGGGATTTATCTGTCAAGGTTGT 1494
Db 1563 GAACGGGCTGGCACCGCTGCGCGCGGCACCTATCGGACGCTGCGCGGCGACCTTC 1622
QY 1495 GGGGGAAACGGATTTCCGTTTCAATGAGTGTGCTTCTGTTCAAGCTTGGGCTGGA 1554
Db 1623 GCGGCTTCGACTGACCGTGGAGAGCGGACCGGACCGACCGGTCGACCGTTCGCGCG 1682
QY 1555 GCGTGTCTGTTGGCAGTACAGCACATCCGCTT- - - - -CAGCGCG 1596
Db 1583 GTGCTGGTCCGGGACAGTTCGCGTCTGGTGTACCGGGCGCGGTGGACACCGAGCC 1742
QY 1597 CAAATCGATTCGGTTCCTCAAAATATGGGATTCGGGTAATGTGTACAGTTCGACGG 1656
Db 1743 CCGATCGGGGCTGGGCGGCTGTCGACCGGGCGGCGCACCGCTCAACCGTTCGAGGCG 1802
QY 1657 AAGGTTTGGGACGACGAGGACCGTGAATTTGGGAGTGAAGAGTGTGAAA 1716
Db 1803 ACGGCTTCGG- - -CTCGCGGGAACCGTGCAGATCGGCGGAGTCCCGGACCGCTCAG 1859
QY 1717 TCTTGGACATCCAAATCGGATTAAGTGTAGCTTC- - -CAACATGGCGCGCGGCTGACC 1773
Db 1860 CAGTGGACGCGGACCGTATCACCGCACCGTCCGGTGGGTTCCACCGGGGCGCTC 1919
QY 1774 GATGTGAAGTACCGCGGGTGGAGTTTCAGCAATCTGTATTTTACATATTTTGTAGT 1833
Db 1920 CAGTGTACCGTGGCAACGGCTCCGGCAACGAGACAGGATACCGGATCACACCGCTACC 1979
QY 1834 GGAACGACAGATTCGGTGTGTTTACTGTGAAGTGGCTCTCGACACCACTGGGGAT 1893
Db 1980 GGAACCGGTCCCGGTGAGTTTACCGTTTCAAGAACCGCGGACCGCGCGCGGGAG 2039

QY 1894 AAGATTACTGACGGGCAACATACCGGAATTGGGAATTGGAGCAGGATACGACCGA 1953
Db 2040 TCGTCTTACTGACCGGTGAGTCTCGCGAGTTGGGCACTGG- - - - -TCGACCGAC 2090
QY 1954 GCGTTAAAGATGCGCAAGGGCCCTGCTCGGCGCCCAATTATCCGGATTGTTTTATGA 2013
Db 2091 CCGAGCAGACCGCGGAGACGCTGCTCGGGTCCGAGAGTCCCGGGCGTCTCTGTC 2150
QY 2014 TTCAAGCTTCCAGCAGGAAAGACGATTCAATTCAGTTTCTTCAAGCGTGGGATGA 2073
Db 2151 GCGACCTGCGCGCGCGCGCTCGAGTTCAGTTTCGTTCAAGTTCGCGCGGCGACGCG 2210
QY 2074 ACGATTCAATGGAGAAATGTTTGAACACCGTGGCCACACTCCACGCGTGAACGGT 2133
Db 2211 ACGTGTACTGGGAGGGTGGTCCACACCGGTACACCGTCCCGCGCGCGCGACCGGC 2270
QY 2134 AACATTACTGTACGTGGCAAACT 2158
Db 2271 ACGACCGCTCACCCTGGCAGCGCT 2295

RESULT 10
US-07-927-316A-1
; Sequence 1, Application US/07927316A
; Patent No. 5409824
; GENERAL INFORMATION:
; APPLICANT: SCHMID, Gerhard
; TITLE OF INVENTION: Gamma-CGTase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Collard & Roe, P.C.
; STREET: 1077 No. 5409824thern Boulevard
; CITY: Roslyn
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11576
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/927,316A
; FILING DATE: 19920923
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP91/00560 A2
; FILING DATE: 22-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Collard, Allison C.
; REGISTRATION NUMBER: 22,532
; REFERENCE/DOCKET NUMBER: SCHMID-PCT
; ATTORNEY/AGENT INFORMATION:
; NAME: Freedman, Edward R.
; REGISTRATION NUMBER: 26,048
; REFERENCE/DOCKET NUMBER: SCHMID-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 365-9802
; TELEFAX: (516) 365-9805
; TELEX: 261176 CRG(UR)
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 2100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-07-927-316A-1

Query Match 9.6%; Score 208.2; DB 1; Length 2100;
Best Local Similarity 47.2%; Pred. No. 2.4e-57;
Matches 965; Conservative 0; Mismatches 1008; Indels 72; Gaps 8;

APPLICANT: STALKER, DAVID
TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: CALGENE, INC.
STREET: 1920 FIFTH STREET
CITY: DAVIS
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1 (a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,202
FILING DATE: 6-JUNE-95
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/016,881
FILING DATE: 11-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 93-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 915-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1988 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO.
US-08-469-202-26

Query Match 3.4%; Score 74; DB 1; Length 1988;
Best Local Similarity 55.3%; Pred. No. 1.7e-13;
Matches 168; Conservative 0; Mismatches 130; Indels 6; Gaps 1;
QY 223 AAGTGGAAAATGTATTGGGGCGGGGATCTGGAGGGGGTTGTCACAAAACCTTCCTTATCTT 282
DB 231 AATTTAAAAAATACTAGTGGAGGAGATCTCGGGGGTTGATTAATAAATACCTTATTTA 290
QY 283 AAACAGCTGGCGGTAAACGAACTCTGGTTGTCGCCGGTTTGGACAACTCTGGATACATG 342
DB 291 AAATCACTGGTGTACTTCAATCTGGATTCTCCCCCAATCGAATAATGTAATAACT 350
QY 343 GCGGGCAGCGATAACAGGCTATACGGATACCTGCGCGGATTTTAAACAGATTGAG 402
DB 351 GATGCTGTGGCAATACCTGATATCATGTTATGGGAGAGATTATTTTCGATAGAT 410
QY 403 GAACATTTCCGGGA-----ATTGGACCAATTTGACAGGTTGGTCAATGATGCTCACAA 456
DB 411 GAACATTTGGCAATCTCGATGATTTTCAAGAACTGACTAGTTTGAATGATAGTCTGAT 470
QY 457 AACGGAATCAAGGTGATTCGACTTTTGCCCAATCATCTGACTCTTTTAAAGGCAAC 516
DB 471 TATAATAGAAAGTGTCTTGATTATGCCCTTAATCATTCGATCGAATGATGAAAT 530
QY 517 GATT 520
DB 531 GATT 534

RESULT 12
US-08-484-434C-33
Sequence 33, Application US/08484434C
Patent No. 5969214
GENERAL INFORMATION:
APPLICANT: STALKER, DAVID
TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene LLC
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC
OPERATING SYSTEM: Windows NT 4.0; MS Word for Windows 7.0a
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,434C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/016,881
FILING DATE: 11-FEB-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Schwedler, Carl J.
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: 15593/01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 530-792-2463
TELEFAX: 530-792-2463
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1988 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
US-08-484-434C-33

Query Match 3.4%; Score 74; DB 2; Length 1988;
Best Local Similarity 55.3%; Pred. No. 1.7e-13;
Matches 168; Conservative 0; Mismatches 130; Indels 6; Gaps 1;
QY 223 AAGTGGAAAATGTATTGGGGCGGGGATCTGGAGGGGGTTGTCACAAAACCTTCCTTATCTT 282
DB 231 AATTTAAAAAATACTAGTGGAGGAGATCTCGGGGGTTGATTAATAAATACCTTATTTA 290
QY 283 AAACAGCTGGCGGTAAACGAACTCTGGTTGTCGCCGGTTTGGACAACTCTGGATACATG 342
DB 291 AAATCACTGGTGTACTTCAATCTGGATTCTCCCCCAATCGAATAATGTAATAACT 350
QY 343 GCGGGCAGCGATAACAGGCTATACGGATACCTGCGCGGATTTTAAACAGATTGAG 402
DB 351 GATGCTGTGGCAATACCTGATATCATGTTATGGGAGAGATTATTTTCGATAGAT 410
QY 403 GAACATTTCCGGGA-----ATTGGACCAATTTGACAGGTTGGTCAATGATGCTCACAA 456
DB 411 GAACATTTGGCAATCTCGATGATTTTCAAGAACTGACTAGTTTGAATGATAGTCTGAT 470
QY 457 AACGGAATCAAGGTGATTCGACTTTTGCCCAATCATCTGACTCTTTTAAAGGCAAC 516
DB 471 TATAATAGAAAGTGTCTTGATTATGCCCTTAATCATTCGATCGAATGATGAAAT 530
QY 517 GATT 520
DB 531 GATT 534

RESULT 13

US-09-384-361-33
; Sequence 33, Application US/09384361
; Patent No. 6538181
; GENERAL INFORMATION:
; APPLICANT: STALKER, DAVID
; TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene LLC
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC
; OPERATING SYSTEM: Windows NT 4.0; MS Word for Windows 7.0a
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/384,361
; FILING DATE: 26-August 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/484,434
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/016,881
; FILING DATE: 11-FEB-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwedler, Carl J.
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: 15593/03/US
; ATTORNEY/AGENT INFORMATION:
; NAME: Eagle, Alissa M.
; REGISTRATION NUMBER: 37,126
; REFERENCE/DOCKET NUMBER: 15593/03/US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 530-792-2265
; TELEFAX: 530-792-2463
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1988 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; US-09-384-361-33

Query Match 3.4%; Score 74; DB 4; Length 1988;
Best Local Similarity 55.3%; Pred. No. 1.7e-13;
Matches 168; Conservative 0; Mismatches 130; Indels 6; Gaps 1;

QY	223	AAGTGGAAATGATTTGGGGCGGAGTCGGAGGGGTTCTGCAAAAACCTCTTACTT	282
Db	231	AAATTTAAAAAATAATCTCGGAGGAGATCTCCGGGGTTGATTAATAACTACCTATT	290
QY	283	AAACAGCTGGCGGAACGACAACTCTGTTGCCCGGTTTGACAACTCGGATACACTG	342
Db	291	AAATCAGTTGGTGTACTTCAATCTGGATCTCCCAATCGAATATGTAATAACT	350
QY	343	GGGGACCCGATAACACGGGCTATCAGGATATCGGACGGCGGATTTTAAACAGATTG	402
Db	351	GATGCTGCTGCAATACTCGGATATCATGTTATTTGGGAAGAGATTATTTTCGTATAG	410
QY	403	GAACATTTTCGGGA-----ATTGGACCAATTTGACAGTTGGTGTGATGCTCACC	456
Db	411	GAACATTTTCGGCAATCTCGATGATTTTCAAGAAGAACTGACTAGTTTGATGATGCT	470

QY	457	AACGGAATCAAGGTGATTTGCGACTTTGGCCCAATCTTCGACTTCCTTTTAAGGCA	516
Db	471	TATAATATGAACTGGTCTTGTGATATGCCCCTAATCATTCGAATGCTAATGATGA	530
QY	517	GATT 520	
Db	531	GAAT 534	

RESULT 14

US-08-469-202-25
; Sequence 25, Application US/08469202
; Patent No. 5750875
; GENERAL INFORMATION:
; APPLICANT: STALKER, DAVID
; TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALGENE, INC.
; STREET: 1920 FIFTH STREET
; CITY: DAVIS
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1 (a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,202
; FILING DATE: 6-JUNE-95
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/016,881
; FILING DATE: 11-FEB-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 93-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 916-753-6313
; TELEFAX: 916-753-1510
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; US-08-469-202-25

Query Match 3.4%; Score 74; DB 1; Length 2000;
Best Local Similarity 55.3%; Pred. No. 1.8e-13;
Matches 168; Conservative 0; Mismatches 130; Indels 6; Gaps 1;

QY	223	AAGTGGAAATGATTTGGGGCGGAGTCGGAGGGGTTCTGCAAAAACCTCTTACTT	282
Db	237	AAATTTAAAAAATAATCTCGGAGGAGATCTCCGGGGTTGATTAATAACTACCTATT	296
QY	283	AAACAGCTGGCGGAACGACAACTCTGTTGCCCGGTTTGACAACTCGGATACACTG	342
Db	297	AAATCAGTTGGTGTACTTCAATCTGGATCTCCCAATCGAATATGTAATAACT	356
QY	343	CGGGACCCGATAACACGGGCTATCAGGATATCGGACGGCGGATTTTAAACAGATTG	402

Db 357 GATGCTGCTGCAACTACTGATATCATGTTATTGGGAGAGAGATTATTTTCGTATGAT 416
 Qy 403 GAACATTTCCGGA-----ATTGGACCAATTTGACAGTTGGTCAATGATCTCACCAA 456
 Db 417 GAACATTTTGGCAATCTCGATGATTTCAAGAAGACTGACTAGTTTGATGATAGTCTCTGAT 476
 Qy 457 AACGGAATCAAGGTGATTCGACTTTGCGCCATCATTCGACTCTCTTTTAAGGCAAC 516
 Db 477 TATAATATGAAGTGGTTCCTTGATTATGCCCTAATCATTCGAATGCTAATGATGAAAT 536
 Qy 517 GATT 520
 Db 537 GAAT 540

RESULT 15

US-08-484-434C-32
 ; Sequence 32, Application US/08484434C
 ; Patent No. 5969214
 ; GENERAL INFORMATION:
 ; APPLICANT: STALKER, DAVID
 ; TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
 ; NUMBER OF SEQUENCES: 35
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Calgene LLC
 ; STREET: 1920 Fifth Street
 ; CITY: Davis
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 95616
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC
 ; OPERATING SYSTEM: Windows NT 4.0; MS Word for Windows 7.0a
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,434C
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/016,881
 ; FILING DATE: 11-FEB-1993
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Schwelker, Carl J.
 ; REGISTRATION NUMBER: 36,924
 ; REFERENCE/DOCKET NUMBER: 15593/01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 530-792-2265
 ; TELEFAX: 530-792-2463
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2000 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHEetical: NO
 US-08-484-434C-32

Query Match 3.4%; Score 74; DB 2; Length 2000;
 Best Local Similarity 55.3%; Pred. No. 1.8e-13;
 Matches 168; Conservative 0; Mismatches 130; Indels 6; Gaps 1;
 Qy 223 AAGTGAATATATTGGGGGGGATCTGGAGGGGTTCGTCAAAACTTCTTATCTT 282
 Db 237 AATTAAAAATACTGGAGGAGATCTCCGGGGGTGATTATAAACTACCCATTATTA 296
 Qy 283 AAACAGCTGGGGCTAACCAATCTGTTGTCCTCCGGTTTGGACATCTGGATACACTG 342
 Db 297 AATCACTTGGTGTACTTCAATCTGGATCTCTCCCAATCGATATGATATACT 356
 Qy 343 GCGGGCACCAGTAACACGGGCTATCAGGGATACAGGCGCGGATTTTAAACAGATTGAG 402

Db 357 GATGCTGCTGCAACTACTGATATCATGTTATTGGGAGAGAGATTATTTTCGTATGAT 416
 Qy 403 GAACATTTCCGGA-----ATTGGACCAATTTGACAGTTGGTCAATGATCTCACCAA 456
 Db 417 GAACATTTTGGCAATCTCGATGATTTCAAGAAGACTGACTAGTTTGATGATAGTCTCTGAT 476
 Qy 457 AACGGAATCAAGGTGATTCGACTTTGCGCCATCATTCGACTCTCTTTTAAGGCAAC 516
 Db 477 TATAATATGAAGTGGTTCCTTGATTATGCCCTAATCATTCGAATGCTAATGATGAAAT 536
 Qy 517 GATT 520
 Db 537 GAAT 540

Search completed: September 17, 2003, 10:52:28
 Job time : 125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 09:59:02 ; Search time 378 Seconds
(without alignments)
14056.674 Million cell updates/sec

Title: US-10-069-908-1
Perfect score: 2160
Sequence: 1 atgaaagaacacgtttc.....ctgttcgtggcaaaactag 2160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1660708 seqs, 1229959015 residues

Total number of hits satisfying chosen parameters: 3321416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/1/pubna/US00_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	2160	100.0	2160	14	US-10-234-266-1
2	234.6	10.9	2631	12	US-10-332-937-1
3	233.2	10.8	2530	14	US-10-188-066-1
4	58.4	2.7	2790	14	US-10-081-872-195
5	54	2.5	5418	14	US-10-156-761-5959
6	54	2.5	9025608	14	US-10-156-761-1
7	49	2.3	1488	14	US-10-213-990-41
8	47.8	2.2	3285	12	US-10-228-083-46
9	46.8	2.2	2863	12	US-10-050-763-2
10	46.4	2.1	1683	14	US-09-974-300-655
11	46.2	2.1	3393	14	US-10-128-714-203
12	45.2	2.1	1509	14	US-10-081-872-107
13	43.8	2.0	1893	14	US-10-213-990-44
14	43.2	2.0	30246	8	US-08-781-986A-56
15	42.2	2.0	1572	14	US-10-081-872-185
16	40.6	1.9	1596	14	US-10-081-872-183

17	40.6	1.9	1596	14	US-10-081-872-191	Sequence 191, App
18	39.8	1.8	4888	10	US-09-070-927A-307	Sequence 307, App
19	38.8	1.8	1945	10	US-09-938-842A-5232	Sequence 5232, App
20	37.4	1.7	1596	14	US-10-081-872-157	Sequence 157, App
21	37.4	1.7	1689	10	US-09-974-300-657	Sequence 657, App
22	37.4	1.7	2489	10	US-09-885-329-1	Sequence 1, Appli
23	37.2	1.7	2250	14	US-10-213-990-43	Sequence 43, Appli
24	37	1.7	6142	13	US-10-014-436-1	Sequence 1, Appli
25	36.8	1.7	3309400	10	US-09-738-626-1	Sequence 1, Appli
26	36.6	1.7	1703	10	US-09-974-300-644	Sequence 644, App
27	36.4	1.7	1704	14	US-10-156-761-6284	Sequence 6284, App
28	36.4	1.7	4190	12	US-10-128-587A-28	Sequence 28, Appli
29	36.4	1.7	4190	14	US-10-128-590-28	Sequence 28, Appli
30	36.2	1.7	582	10	US-09-974-300-5148	Sequence 5148, App
31	36.2	1.7	1638	9	US-09-815-242-4455	Sequence 4455, App
32	36.2	1.7	1650	9	US-08-815-242-8494	Sequence 8494, App
33	35.8	1.7	1431	14	US-10-081-872-171	Sequence 171, App
34	35.8	1.7	1587	14	US-10-081-872-159	Sequence 159, App
35	35.8	1.7	1596	14	US-10-081-872-167	Sequence 167, App
36	35.8	1.7	1596	14	US-10-081-872-189	Sequence 189, App
37	35.8	1.7	1623	14	US-10-081-872-161	Sequence 161, App
38	35.4	1.6	2437	14	US-10-238-091-6	Sequence 6, Appli
39	35.4	1.6	2997	14	US-10-238-091-2	Sequence 1, Appli
40	35.4	1.6	2997	14	US-10-238-091-2	Sequence 2, Appli
41	35.2	1.6	405	10	US-09-738-626-981	Sequence 981, App
42	35.2	1.6	1881	14	US-10-081-872-131	Sequence 131, App
43	35	1.6	1770	14	US-10-081-872-123	Sequence 123, App
44	34.8	1.6	472	13	US-10-027-632-293710	Sequence 293710, App
45	34.8	1.6	569	13	US-10-027-632-39673	Sequence 39673, A

ALIGNMENTS

RESULT 1
US-10-234-266-1
; Sequence 1, Application US/10234266
; Publication No. US20030059902A1
; GENERAL INFORMATION:
; APPLICANT: Cherry, Joel
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Beier, Lars
; APPLICANT: Frandsen, Torben
; APPLICANT: Schafer, Thomas
; TITLE OF INVENTION: Amyolytic Enzyme Variants
; FILE REFERENCE: 5241.204-US
; CURRENT APPLICATION NUMBER: US/10/234,266
; CURRENT FILING DATE: 2002-09-04
; PRIOR FILING DATE: US/09/645,707
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2157)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (100)..()

Query Match 100.0%; Score 2160; DB 14; Length 2160;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGAAACGTTTCTTTATTTGTGGACTGATGCTCTCATCGTCTTCTTC 60
|||||

1 ATGAAAGAAAGCGCTTCTTTATTTGTGGAGCTGATGCTCTCATCGGTCTTCGTTC 60
61 AGCGGTTCCTTCGTCAATCAAAACGCGCTGAAGCAGAGTTCGCGAAGCGTCAAA 120
61 AGCGGTTCCTTCGTCAATCAAAACGCGCTGAAGCAGAGTTCGCGAAGCGTCAAA 120
121 GGGACGTGATTACAGATTATCATGACCGGTTTACGATGGGACACGACCAAC 180
121 GGGACGTGATTACAGATTATCATGACCGGTTTACGATGGGACACGACCAAC 180
181 AATCTGCGAAAGATTATGAGCTTACGATCGGACCAATTCGAAAGTGGAAATGTATGG 240
181 AATCTGCGAAAGATTATGAGCTTACGATCGGACCAATTCGAAAGTGGAAATGTATGG 240
241 GCGGGGATCTGGAGGGGTTGTGCAAACTTCTTATCTTAAACAGCTGGGCTGAACG 300
241 GCGGGGATCTGGAGGGGTTGTGCAAACTTCTTATCTTAAACAGCTGGGCTGAACG 300
301 ACAATCTGTTGTCCCGGTTTGGCAATCTGGATACACTGGCGGACCGGATAACG 360
301 ACAATCTGTTGTCCCGGTTTGGCAATCTGGATACACTGGCGGACCGGATAACG 360
361 GCGTATCGGATCTGAGCGCGGATTTTAAACAGATTGAGGAACATTTGGGAATTGG 420
361 GCGTATCGGATCTGAGCGCGGATTTTAAACAGATTGAGGAACATTTGGGAATTGG 420
421 ACCAATTTGACAGCTTGTGCAATGATGCTCACAAACGGAATCAAGTGATTTGCGAC 480
421 ACCAATTTGACAGCTTGTGCAATGATGCTCACAAACGGAATCAAGTGATTTGCGAC 480
481 TTTGTGCCAATCATTCGATCTCCTTTTAAAGCAACAGATTTCACCTTTCCGGAAGGCGC 540
481 TTTGTGCCAATCATTCGATCTCCTTTTAAAGCAACAGATTTCACCTTTCCGGAAGGCGC 540
541 GCGCTCAACAATGGAACCTATATGGCAATTTATTTGATGACGCAACAAAGAGGTAC 600
541 GCGCTCAACAATGGAACCTATATGGCAATTTATTTGATGACGCAACAAAGAGGTAC 600
601 TTCCACCAATATGGGACATCAGCAACTGGGACGCGGTACGAGGCGCAATGGAAAC 660
601 TTCCACCAATATGGGACATCAGCAACTGGGACGCGGTACGAGGCGCAATGGAAAC 660
661 TTACGCGATCAGCGGTTCTCGCTTCCGATTTTTCGAGGAATGSCACGATTGCT 720
661 TTACGCGATCAGCGGTTCTCGCTTCCGATTTTTCGAGGAATGSCACGATTGCT 720
721 CAATACCTGACCGATGCGGGGTTCAATTGGTAGCAGATGGAGCGGATTTGCGGATT 780
721 CAATACCTGACCGATGCGGGGTTCAATTGGTAGCAGATGGAGCGGATTTGCGGATT 780
781 GATGCGGTGACGATTTTAAATGGGTTCTCCAAATCGTTGGCCGATTAACGTACCA 840
781 GATGCGGTGACGATTTTAAATGGGTTCTCCAAATCGTTGGCCGATTAACGTACCA 840
841 AAGAAAGACATTTCTGTTGGGGGATGATACGAGATGACCCCGGACACGCAATCAT 900
841 AAGAAAGACATTTCTGTTGGGGGATGATACGAGATGACCCCGGACACGCAATCAT 900
901 CTGGAAGAGTCGCTAGCGCAACACAGCGGTGCAATGTGCTGATTTGATCTCAAC 960
901 CTGGAAGAGTCGCTAGCGCAACACAGCGGTGCAATGTGCTGATTTGATCTCAAC 960
961 ACSTGATTCGAAATGCTTCGGACATTTACGCAACGATGATACGATCTTAAACATATG 1020
961 ACSTGATTCGAAATGCTTCGGACATTTACGCAACGATGATACGATCTTAAACATATG 1020
1021 GTGAACCAACCGGGAACGAGTACAAATCAAGAAATCTAATCAATTTGATTAAC 1080
1021 GTGAACCAACCGGGAACGAGTACAAATCAAGAAATCTAATCAATTTGATTAAC 1080
1081 CATGATATGTCAGATTTCTTTCGTTAAATTCGAACAGGCGAATTTGCAACGGCGCTT 1140
1081 CATGATATGTCAGATTTCTTTCGTTAAATTCGAACAGGCGAATTTGCAACGGCGCTT 1140

RESULT 2

US-10-332-937-1

; Sequence 1, Application US/10332937

1141 GCTTTCAATCTCACTTCGCGGGGTACGCCCTCCATCTATTATGGAACCGAACATACATG 1200
1141 GCTTTCAATCTCACTTCGCGGGGTACGCCCTCCATCTATTATGGAACCGAACATACATG 1200
1201 GCAGGCGGCAATGACCCGCTACAACCGGGGGATGATGCGCGGTTTGATACGACAAACACC 1260
1201 GCAGGCGGCAATGACCCGCTACAACCGGGGGATGATGCGCGGTTTGATACGACAAACACC 1260
1261 GCTTTAAAGAGGTGTCACTCTGCGGGGTTGCCAGAGAACATCGCGGATTCAGTAC 1320
1261 GCTTTAAAGAGGTGTCACTCTGCGGGGTTGCCAGAGAACATCGCGGATTCAGTAC 1320
1321 GGCAACACACCGACCGTGGATCAACAATGATGTTTACATTTATGAAGGAAATTTTC 1380
1321 GGCAACACACCGACCGTGGATCAACAATGATGTTTACATTTATGAAGGAAATTTTC 1380
1381 AAGATGTCGTGTTGGTGGCCATCAATCGAAACACGCAATCTCTTATTCGATTTCCGGT 1440
1381 AAGATGTCGTGTTGGTGGCCATCAATCGAAACACGCAATCTCTTATTCGATTTCCGGT 1440
1441 TTGACAGCGGCTTCGCAATGAGCTATGCGGATTAATCTGTCAGGGCTGTTGGGGGG 1500
1441 TTGACAGCGGCTTCGCAATGAGCTATGCGGATTAATCTGTCAGGGCTGTTGGGGGG 1500
1501 AACGGATTTCCGTTTCCAAATGGAAGTGTGCTTCGTTCAAGCTTGGCGCTGGAGCGGTG 1560
1501 AACGGATTTCCGTTTCCAAATGGAAGTGTGCTTCGTTCAAGCTTGGCGCTGGAGCGGTG 1560
1561 TCTGTTGGCAGTACAGCACATCGCTTCAGCGCGCAAAATCGGATCGGTGCTCCAAAT 1620
1561 TCTGTTGGCAGTACAGCACATCGCTTCAGCGCGCAAAATCGGATCGGTGCTCCAAAT 1620
1621 ATGGGATTTCCGGTAAATGTTGTCAGCTGACGCGGAAAGGTTTGGGACGACGAGGGA 1680
1621 ATGGGATTTCCGGTAAATGTTGTCAGCTGACGCGGAAAGGTTTGGGACGACGAGGGA 1680
1681 ACCGTGACATTTGGCGGAGTACAGCGACTGTGGAATCTTGCAATCCATCGGATGGA 1740
1681 ACCGTGACATTTGGCGGAGTACAGCGACTGTGGAATCTTGGAATCCATCGGATGGA 1740
1741 GTGTACCTTCCCAACATGGCGCGGCTGACCGATGTGAAGTCAAGCGCGGTGGAGTT 1800
1741 GTGTACCTTCCCAACATGGCGCGGCTGACCGATGTGAAGTCAAGCGCGGTGGAGTT 1800
1801 TCCAGCAATCTGTATTTCTTACAATATTTTGAAGTGAAGTCAAGCGCGGTGGAGTT 1860
1801 TCCAGCAATCTGTATTTCTTACAATATTTTGAAGTGAAGTCAAGCGCGGTGGAGTT 1860
1861 GTGAAAGTGCCTCCGACCACTCGGGGATAGATTTACCTGACGCGCAACATACCG 1920
1861 GTGAAAGTGCCTCCGACCACTCGGGGATAGATTTACCTGACGCGCAACATACCG 1920
1921 GAATTTGGGAAATTTGGAGCACGGAATACGCGGAGCGGTTAAACATTCGCAAGGCGCGCTG 1980
1921 GAATTTGGGAAATTTGGAGCACGGAATACGCGGAGCGGTTAAACATTCGCAAGGCGCGCTG 1980
1981 CTCGCGCCCAATTTCCGATTTGTTTATGATTTTCAAGCTTCCAGCAGGAAGAGATT 2040
1981 CTCGCGCCCAATTTCCGATTTGTTTATGATTTTCAAGCTTCCAGCAGGAAGAGATT 2040
2041 CAATTCAGTTCTTCAACAGCGTGGGATGGAATCAATGGGAGATTTGTTCCGAAC 2100
2041 CAATTCAGTTCTTCAACAGCGTGGGATGGAATCAATGGGAGATTTGTTCCGAAC 2100
2101 CAGTGGCACAATCTCCACGCGGTGCAACCGGTAACTTACTGTTAGTGGGAAAATAG 2160
2101 CAGTGGCACAATCTCCACGCGGTGCAACCGGTAACTTACTGTTAGTGGGAAAATAG 2160

; Publication No. US20030170696A1
 ; GENERAL INFORMATION:
 ; APPLICANT: No. US20030170696A1ozymes
 ; APPLICANT: Jorgensen, Per Lina
 ; APPLICANT: Fuglsang, Claus Crone
 ; TITLE OF INVENTION: Cgt'ase and DNA sequence encoding same
 ; FILE REFERENCE: 10037.204-WO
 ; CURRENT APPLICATION NUMBER: US/10/332,937
 ; CURRENT FILING DATE: 2003-01-14
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 2631
 ; TYPE: DNA
 ; ORGANISM: Bacillus agardhaerans
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: (103)..()
 ; OTHER INFORMATION:
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2139)
 ; OTHER INFORMATION:
 ; US-10-332-937-1

Query Match 10.9%; Score 234.6; DB 12; Length 2631;
 Best Local Similarity 48.5%; Pred. No. 9.6e-66;
 Matches 991; Conservative 0; Mismatches 989; Indels 63; Gaps 10;

QY	124	GACGTGATTACCAAGATTATCATGACCGGTTTTACGATGGGACACGACGACAAACAAT	183
DB	145	GATGTCATTTATCAAAATTGACAGATAGTTTTACGATGGTGTATCAAGATTAACACCCA	204
QY	184	CTGCGCAAAAGTTATGACTTTTACGATCGACCAATTCGAGTGGGAAATGNTTGGGGC	243
DB	205	TCAGAGAACTTTAT---TCGGAAGATTGTAACAACTTAAGAAATATTTGGTGGAGAT	261
QY	244	GGGATCTGGAGGGGTTCTCAAAATCTCTTATCTTAAACAGCTGGGCGTACGACA	303
DB	262	TGGCAAGGATAATAGATATAATAGATGATGATGTTATCAAGACATGGGTGACGGCA	321
QY	304	ATCTGGTGTCCCGGTTTTGGCAATCTGGATACATG-----CGGGACCGATAAC	357
DB	322	CTATGGATCTCACCCCGGTTGAAATATTTTGAACGATGATGATGATGATGATGATGAT	381
QY	358	ACGGGCTACCGGATCTGACGCGGATTTTAAACAGATTGAGCAATCTGGGAAAT	417
DB	382	ACTTCTTACGCTTATGGGACGAGATTATAGAAACGAAACCCCTTTTGGAGC	441
QY	418	TGGACCAATTGACAGCTTGGTCAATGATGCTCACAAACGGAATCAAGGTGATGTC	477
DB	442	ACAGAAGATTTTGAAGGTTAATAGAAACTGCACATAGTCACGATATTAATATGTTATT	501
QY	478	GACTTGGCCCATCTTCTGACTCTTTTAAAGGCAACGATTCACCTTTCCGGAGGC	537
DB	502	GATTTAGCTCCATACCATATCACCTGAGATTTTGAATATCTCACTATCCGAAAT	561
QY	538	GGCGCCCTCTACAAATGAAACCTATATGGGCAATTTTGGATCAGCACAACAAAGG	597
DB	562	GGTGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	621
QY	598	TACTTCCACCATTAAGGGGACATCAGCAATCTGGGACGACCGGTACGAGGCGAATGAAA	657
DB	622	TTATATAAGGGGAAACAGATTTCTCTAACTACGAAGAT-----GAGATTTATAGA	672
QY	658	AACCTTCAGGATCCAGCGGTTCTCTGCTTGGGATTTTGGCGAGGAATCGCAGATT	717
DB	673	AATTTGTTGACTTAGCTAGTTTTTAATCATATCATCATCATCATCATCATCATCAT	723
QY	718	GCTCAATACCTGACCGATGGGGGTTCAATTTGGTAGCATGATGAGCGGATGTTGGG	777
DB	724	-----TATTTAGAGATGAGTGAAATAATGTTAGATTAGATTAGATTAGATTAGATTAG	777

QY	778	ATTGATGCGGTGAGACATTTTAATTCGGGGTCTCCAAATCGTTGGCCGATAAACGTGAC	837
DB	778	ATCGATGCTGTAGCTCAATGCCACCAGGTTCGAAAAAAGCTTACATGGATACATATAT	837
QY	838	CAAAAGAAAGACATTTCTCTGGTGGGGGATGATGAGGATGATGATGATGATGATGATGAT	897
DB	838	GACCAAGAGCGGTTATTTCTTTTGGAGAAATGTTTACTGGACCTTCTGGAATGAGGAT	897
QY	898	CATCTGAAAAGGTCGGGTACGCAACAAACAGCGGTGTCAATGTGCTGATTTTGTATCTC	957
DB	898	TACACTAAA-----TTTGCATAATATAGTGGCATGAGTATTAGATTTCCGCTTT	948
QY	958	AACACGGTGAATCGAAATGTTTGGGCAATTTAGCGAAACGATGATGATGATGATGATGAT	1017
DB	949	GCTCAAACTACACGAAATGTCTCGGTAAACAATAATGAACGATGATGATGATGATGATGAT	1008
QY	1018	ATGGTCAACCAAAACGGGACGAGTACAAATACAAAGAAATCTAAATCAATTTATCGAT	1077
DB	1009	ATGCTAACAGACACAGAGATGACTATGATCGTCTCAAGATCAAGTTACTTTTCTTGAT	1068
QY	1078	AACATGATATGTCAGATTTCTTTTGGTAAATTCGAACAAGCGGAATTTGCACAGGGG	1137
DB	1069	AATCATGACATGATGATGATTTTACGAATGTTGGTGAATCAACACGACACACAGATATTGA	1128
QY	1138	CTTGTCTTCAATCTCACTTCGCGGGTACGCCCTCATCTTATATGAAACCGAACATAC	1197
DB	1129	TTAGCTTAAATGCTTACATCTCTGGGGTGCCTACTATTTATTTATGAAACAGACATAC	1188
QY	1198	ATGCAAGGTTGAGGATCCAGGAAGCAGGGGAATGATGGCATCTTTTGTGAAAAATACA	1248
DB	1258	ACCGCTTTAAAGAGGTCTCACTCTGGCGGGTTCGCGACGACAAATGCGGATCCAG	1317
QY	1249	GATGCTTAAAGCTAAATTCAAAAATTAGACCGCTTAAGAAAAAGTAACTCTGATATGA	1308
DB	1318	TACGGACACACACACCGGTTGGATCAACATGATGTTTACATTTATGAACGGAATTT	1377
QY	1309	TATGACACACACACGACGTTGGTAAACGATGATGATGATGATGATGATGATGATGATGAT	1368
DB	1378	TTCAAGATGCTGTTGGTGGCATCAATCGAAACACGCAATCTCTTATTCGATTTCC	1437
QY	1369	GGCGAGAACTATGCTTTAATCGCATTAATAGAGCTTAAATACCTCTATAATATCCAA	1428
DB	1438	GTTTTCGACAGCGCTTCCCAAATGCGAGTATGCGGATTTCTGTCAAGGCTGTTGGGG	1497
QY	1429	GGTTTACACACAGATGCGCATCAATCATATGATGATGATGATGATGATGATGATGATGAT	1488
DB	1498	GGGAACGGGATTTCCGT---TTCCAATGGAAGTGTGGCTTCGTTCAAGCTTGGCGCTGA	1554
QY	1489	GGGCAATCAATTTGTTGATAACAAAGGGGAGTTAATGAATTTCAAAATGCTCTCTGGA	1548
DB	1555	CCGCTGCTGTTTGGCAGTACAGACATCCGCTTCAG---CGCGCAAAATCGATCGGTT	1611
QY	1549	GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1608
DB	1612	GCTCCAAATATGGAATTCGGGTAAATGTTGTCACGATCGAGGAAAGGTTTGGGAGG	1671
QY	1609	GGCCCAATTAATTTGGTGAGGACGACGATGTTTACAAATGATGAGGAGGTTTCCGTTCT	1668
DB	1672	ACGAGGAAACGTCAGCAATTTGGCGGAGTACAGCGATGTCGAAATCTGACATCCAAAT	1731
QY	1669	TCGAGGAGGACTGTTCACTTTGGATCCACTTCAGCAGAAATCTTTCTTGGAAATGATCG	1728
DB	1732	CGGATTAAGTGTAGTTTCCCAACATCGCGCGGCTGACCGATGTA---AAGTCAAC	1788
QY	1729	ATCATTAACCTTAACCTGTGCGCAACATGAAGCGGATACCATGATATCATCTGTTGAACA	1788
DB	1789	CGCGGTGAGTTTCCCAATCTGATTTCTTCAATAATTTTGAAGTGAACGACGATCG	1848
QY	1789	GAAGATGAACAAAGTAAGTAATGCTATGAATTCGAAGTTCTTAACGGCGGATCAAGTACA	1848
DB	1849	GTTGTGTTACTGTGAAAGATGGCCCTCCGACCAACCTGGGGGATAGATTTACCTGACG	1908

Db	1849	GTTCGTTTATCATAGACAATG---CAGAAACCAAGCTGGGTGAAACGTTTTCCTTGTA	1905
Qy	1909	GGCAACATACCGGAATTGGGGAATTGAGCAGCGGATACGAGCGGAGCGGTTAAACAATGCG	1968
Db	1906	GGTAACGTTTCATGANTTAGGAANTGGGATCCAGAACAAATCAGTTGGGAGATTTTCAAT	1965
Qy	1969	CAAGGGCCCGTGTCCGCGCCCAATTAFCGCGATGCGTTTATGTTATTCAGCGTTCCAGCA	2028
Db	1966	CAA-----ATTGTATACCAATATCCAACATGGTATTATGATGTGAATGTTCTGCA	2016
Qy	2029	GGAAAGACGATTCAATTCAAGTCTTCATCAGACGCTGGCGATGGAACGATTCAATGGGAG	2088
Db	2017	AATACAGACGTTAGAATTCAAGTTATTAAATAGATCAAGATAATAACGTCATTTGGCGAG	2076
Qy	2089	AATGGTTTCGAACCAACGCTGGCCACAACTCCCGCGGTGCAACCGGTACATTTACTGTTACG	2148
Db	2077	AGTGGCGTAAATCAAACTTCTTCACAGAAAGTGGAACTGGTATTATAGAGATTGAT	2136
Qy	2149	TGG 2151	
Db	2137	TGG 2139	

RESULT 3

```

US-10-188-066-1
; Sequence 1, Application US/10188066
; Publication No. US20030114417A1
; GENERAL INFORMATION:
; APPLICANT: TAKADA, MASAYASU
; APPLICANT: IDE, TAKAHIRO
; APPLICANT: NAKAGAWA, YOSHINORI
; APPLICANT: YAMAMOTO, TAKESHI
; APPLICANT: YAMAMOTO, MIKIO
; TITLE OF INVENTION: PRODUCING Y-CYCLODEXTRIN AND USE THEREOF
; TITLE OF INVENTION: PRODUCING Y-CYCLODEXTRIN AND USE THEREOF
; FILE REFERENCE: 12637/P8017US0
; CURRENT APPLICATION NUMBER: US/10/188,066
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: JP 2001-211340
; PRIOR FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 1
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Bacillus clarkii
; FEATURE:
; OTHER INFORMATION: 7364
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)..(2426)
US-10-188-066-1

```

QY	361	GGCTATCAGGATACCTGGACCGCCGATTTTAAACAGATGTAGGAACATTTTCGGGAATGG	420
Db	681	TCCTTATCAGGGTATTTGGGCCCGAGACTTTAAAGAAGCAAAACCTTTCTTCGGAGATTTT	740
QY	421	ACCACATTTTGCACAGTTGGTCAATGATGCTCACAAACAGGAATCAAGTGAATCTCGAC	480
Db	741	GATGATTTTTCCGAGCTAATTCGAAACAGCTCATGCAATGACATAAAAGTAGTTATTTGAT	800
QY	481	TTTGTGCCAATCATCTGCATCTCTTTTAAAGCAAAACGATTCACCTTTTCGGGAAGCGGC	540
Db	801	TTTGTACCTTAACCATCTTCCCTGTAGACATCCAGGAT-----GGT	842
QY	541	GCCTCTCAACAATGGAACCTATATGGGCATTTATTTTGTACGCAACAAAGGCTAC	600
Db	843	GCATTTGTATGACACGGTACATTTACTTGGGCACATTTCAACGGATCAACCAATTTATTTT	902
QY	601	TTCCACCAATAATGGGGACATCAGCAACTGGGACACCGGTACGAGCGCAATGGAAAAAC	660
Db	903	TATAACTATGGTGG-----TTCAGACTTCTCAGACTATGAAAT	941
QY	661	TTTACGGATCCAGCCGGTTTTCTCGCTTCCGATTTGTTCGAGAAAAATGCGACGATTGCT	720
Db	942	AGCATCTATCGAAACTTGTATGATTTTAGCAAGTCTTAACGAGCAACATTCCTTTATTTGAT	1001
QY	721	CAATACCTCAGCGATGGCGGGTCAATTTGGTATGACATGAGCGGATGGTTTTCGGGAT	780
Db	1002	AAATACTTAAAGAAATCTATTCAATTTATGTTTGGATACGGGAATTCAGCGGATTCGCGTG	1061
QY	781	GATCGGTGAGCATTTTAATTCGGGTTCTCCTAAATGTTTGGCGGATTAATCTGTACCA	840
Db	1062	GATCGGTTGCAACATGCTCTTTGGCTGGCAAAAGCATTTATCTCATCTGCTCATGAT	1121
QY	841	AAGAAAGACATTTTCTGGTGGGGGAATGTACGAGAGTACCCCGGAACGCCCAATCAT	900
Db	1122	TACATCCAGTTTTTACCTTTGTGTGATGGTTTACAGAGCAACAGCGAGCAATCAITAC	1181
QY	901	CTGAAAAGGTCGGTAGCCCAACACACGGGTGTCAATGTGTGGATTTTGAATCTCAAC	960
Db	1182	C-----ACCATTTTGTCAACAACAGTGGCATGAGCGCTTGTANTTTTCGTATGCT	1232
QY	961	ACGGTATTCGAAATGTGTTGGGCACATTTACGCAAAACGATACCATCTTAACAATATG	1020
Db	1233	CAGTAGCCGAGATGATTAAGAAATCAAAAGGACGATGATGCAATTTACGACATG	1292
QY	1021	GTGACCAACCGGGAACGAGTCAAAATCAAGAATACTTAATCACTTTATCGATAAC	1080
Db	1293	TTGGCAAGCACTCAATTAGATTTATGACGGCGCGCAAGATCAAGTAACCTTTTATTGATAAT	1352
QY	1081	CATGATATGTCAAGATTTCTTTTCGTAAATTCGACAGGCGGAATTTGACACAGCGCTT	1140
Db	1353	CAATGATCATGCTGCTTTACGGTGGAGGCGGAGTACAGGACACAGGACATCCGACATG	1412
QY	1141	GCTTTCATTTCTCACTTCGGGGGTACGCCCTCCATCTATTATGGAACCGAAACAATACATG	1200
Db	1413	GCAATCTTTTTCACATCAAGAGCGTACCGGCTATTTATTTATGTCAGGAAACTATATG	1472
QY	1201	GCAGCGGCAATGACCCGTACACCGGGGGATGATCCGGCGTTTGTATGACACACACC	1260
Db	1473	ACTGGTAAAGGAGATCCAGAGAACAAAAAATGATGGAGAGCTTTGATCAAAACAACGACA	1532
QY	1261	GCTTTTAAAGAGGTGTCAACTCTGGCGGGTTGGCAGGAACATCGCGCATCCAGTAC	1320
Db	1533	GCCTATCAGTCAATCCAAAGCTTGGCACCGCTCCGACAGAAATTAACGGGTGTATAT	1592
QY	1321	GGCAACACACCCAGCTGTGATCAACAATGATGTTTCAATTTATGAACGGAATTTTTC	1380
Db	1593	GGTTCAACAAAAAGAACGTTGGATTAACGATGATGTGCTCATTTATCAACGATCGTTAAT	1652
QY	1381	AAGGATGTGCTGTGGTGGCCATCAATCGAAACACGAACTCCTTATTTGATTTTCGGT	1440
Db	1653	GGAGATATCTTTTATGTGCGCAATTAATAAATATTAATACAGCTATTACTATTTTCGGT	1712
QY	1441	TTGACAGCGCCTTGCCTAAATGCGAGCTATGCGGATTTATCTGTACGGGTGTGTGGGGGG	1500

```
Db 1713 TTGCTCAGGAAATGCCGCGCAAGTCTATCATGATTTTACAGAGCTATTAGACGGT 1772
Qy 1501 AACGGGATTCGGT---TTCATGGAAGTTCGGTTCGTTACAGCTTCGGCTCGAGCC 1557
Db 1773 CAACTGTAGCAGTAAAGAAATGCTACAGTTGATTCCTTCTGCTAGGACCAAGTGA 1832
Qy 1558 GTGCTGTTTGGCAGTA---CAGCAGATCCGCTTCAGCGCGCAATTCGATCGGTTGCT 1614
Db 1833 GTAAGTGTATGGCAGCATATAAGTGAAGTGTTCGCTCTGTTATTTGGTCAAGTNGC 1892
Qy 1615 CCAAAATATGGGATTCGGGTAATGTGTCAGCATCGAGCGGAAGGTTTGGGACGAGC 1674
Db 1893 CCGCTATGGGAAACCTGGAGATGCTGTGAAGATTAGTGGCAGGATTTGGTCTGAG 1952
Qy 1675 CAGGGAACGTGACATTTGGCGGAGTGA CAGCGA CTGGAATCTTGGACATCCAATCGG 1734
Db 1953 CTTGGCCGCTGTACTTCAGAGATAGCAAAATAGACGTGTTAACCTTGGGATGATGAACG 2012
Qy 1735 ATTGAAGTGTACGTTCCCAACATGCGCGCGCGCTGACCGATGGAAGTCAC---CGCG 1791
Db 2013 ATTGTGATCACACTGCCGGAACCATTAGAGGAAGCGCAAAATCAGTGTACTAATCTCT 2072
Qy 1792 GTGCGAGTTTCCAGCAATCTGTATTTCTTACAAATATTTTGTGAGGAAAGCGGACATCGGTT 1851
Db 2073 GACGGGTGACAGAGTAACGGCTATGATTTTCAGTTGTTGACAGGTAAGCAGGATCTGT 2132
Qy 1852 GTGTTTACTGTGAAGAGTGGCCCTCGACGACACCTTGGGGGATAGATTTACCTGACGGGC 1911
Db 2133 CGTTTCGTTGTGATGTAATGCGCAT---ACCAATTTATGGGGAATGTTTATCTTGTGGA 2189
Qy 1912 AACATACCGGAATGGGGAATTTGAGACAGGATAGCGGAGCGGTTAACATCGCAA 1971
Db 2190 ATGTTCTGAGCTTGGGAATTTGGAACCTGCGGACCAATCGGACCAATGTTTAAATCA 2249
Qy 1972 GGGCCCTGCTCGGCGCCAAATATTCGGGATTTGTTTATGTAATTCAGCGTTCCACAGGA 2031
Db 2250 G-----TGTATTTCCTATCCAACTCGTATACGATGTCAGTGTCCCGCGGAT 2300
Qy 2032 AAGACATTCATTCAGTCTTCATCAGCGTGGGATGGAACGATTCATGGAGAT 2091
Db 2301 ACCCGTTGGAATTTAAGTTTATTTATTTGTCGATGGAATGGAATGTTACTTGGGAAAGC 2360
Qy 2092 GGTTCGAACACGCTGGCCACAACTCCACCGGTGCAACCGGTACATTAATCTGTACGTGG 2151
Db 2361 GGGGGTAATCACAATTATCGTTGTTACTCGGGAACGAGATCTGTTGTTAGTTT 2420
Qy 2152 CAAACTA 2159
Db 2421 CGAAGGTA 2428
```

```
RESULT 4
US-10-081-872-195
; Sequence 195, Application US/10081872
; Publication No. US20030125534A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Keruvu, Jaane S.
; APPLICANT: Slupeska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/081,872
; PRIORITY FILING DATE: 2002-02-21
; PRIORITY APPLICATION NUMBER: US 60/270,495
; PRIORITY FILING DATE: 2001-02-21
; PRIORITY APPLICATION NUMBER: US 60/270,496
```

```
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195
; LENGTH: 2790
; TYPE: DNA
; ORGANISM: Environmental
US-10-081-872-195

Query Match 2.7%; Score 58.4; DB 14; Length 2790;
Best Local Similarity 52.4%; Pred. No. 1.5e-07;
Matches 196; Conservative 0; Mismatches 151; Indels 27; Gaps 2;

Qy 139 ATTATCATTCACCGGTTTTACGATGGGACAGCAGCAACAAATCTCTGCCAAAAGTTAT 198
Db 1 ATGCTGACAGACCGGTTTCTTGTGCGGATACATCAACCAACGACCCCTTACAA----- 53
Qy 199 GGACCTTTACGATCCGACCAATCGAAGTGAAGATGTTTGGGGGGGATCTCGAGGGG 258
Db 54 --CCAGAACACTACGATGCTAAACACGCGGGGAACCTTATCAGGGCGGGGATTTAAAGGA 111
Qy 259 GTTCGTCAAAAACTTCCTTTATCTTAAACAGCTGGGCGTAAACGACATCTGGTTGTCGCCG 318
Db 112 ATCAGCGCAAAATTTGGATTATCTCGATTAAGCTAGGCGTGAACACATCTGGATCAGCCG 171
Qy 319 GTTTTGGACAACTCTGGATACACTGGCGGGGCAACCGATTAACACGGGC----- 363
Db 172 ATCGTGGAAATATCAAGCATGATGTCGGTTATGCAACTCTGAAGGGCATTCATCTAT 231
Qy 364 ---TATCAGAGTACTGAGCGCGGATTTTAAACAGATTGCGGAGACATTCGGGAATTGG 420
Db 232 GCTTACACGGCTACTGGGCAAGCACTTCGGTCCGTTAAACCCACRCTTCGTTCAATG 291
Qy 421 ACCACATTTGACAGTTTGGTCAATGATCTCAACAAACGGAATCAAGGTGATTGTGAC 480
Db 292 GAAGATTTCATACACTGATGATGACCTCCCATGCCCATGAAAAGGCATCAAGATCATGTTGAC 351
Qy 481 TTTGTGCCCATCA 494
Db 352 GTAGTATTAAACCA 365
```

```
RESULT 5
US-10-156-761-5959
; Sequence 5959, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 5959
; LENGTH: 5418
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(5418)
US-10-156-761-5959
```

Query Match 2.5%; Score 54; DB 14; Length 5418;
Best Local Similarity 53.0%; Pred. No. 6.6e-06;
Matches 142; Conservative 0; Mismatches 120; Indels 6; Gaps 1;

Qy 233 TGTATTGGGGGGGATCTGGAGGGGTTCTCAAAAATTCCTTATCTTAAACAGCTGG 292
Db 311 TCTACAGGGGGGACCTCAAGGGCGTCAAGGGCGTCAAGGGCGTCAAGGGCGCTCG 370

Qy 293 GGTAAACAGCAATCTGGTTCTCCCGGTTTGGACAATCTGGATACACTGGGGGACCG 352
Db 371 GCACCACCTCATCTGGCTGGCGCGATCTTCAAGAACCGCCGCTGAGGCGACCGCA 430

Qy 353 ATAAAC-----ACGGGCTACGGATACCTGACGCGCGATTTTAAACAGATTGAGGAAC 406
Db 431 AGGACGCTCGCGCGCTACCGGCTACTGGATCACCAGCTTCAACAGGTTCGACCCGC 490

Qy 407 ATTTCGGGAATTGGACCACTTTCAGACGTTGGTCAATGATGCTACCAAAACGGATCA 466
Db 491 ACTTCGGTACCAACAGGACCTCGAAACCCCTCATCTCCAGGGCGCACGCAAGGGCATGA 550

Qy 467 AGTGATTGTGCACTTTGTGCCCAATCA 494
Db 551 AGGTCTTCTGACGTCATCACCAACCA 578

RESULT 6
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 2.5%; Score 54; DB 14; Length 9025608;
Best Local Similarity 53.0%; Pred. No. 0.0011;
Matches 142; Conservative 0; Mismatches 120; Indels 6; Gaps 1;

Qy 233 TGTATTGGGGGGGATCTGGAGGGGTTCTCAAAAATTCCTTATCTTAAACAGCTGG 292
Db 7202597 TCTACAGGGGGGACCTCAAGGGCGTCAAGGGCGTCAAGGGCGTCAAGGGCGCTCG 7202656

Qy 293 GGTAAACAGCAATCTGGTTCTCCCGGTTTGGACAATCTGGATACACTGGGGGACCG 352
Db 7202657 GCACCACCTCATCTGGCTGGCGCGATCTTCAAGAACCGCCGCTGAGGCGACCGCA 7202716

Qy 353 ATAAAC-----ACGGGCTACGGATACCTGACGCGCGATTTTAAACAGATTGAGGAAC 406
Db 7202717 AGGACGCTCGCGCGCTACACCGGCTACTGGATCACCGGCTCAACCAAGGTTCGACCCGC 7202776

Qy 407 ATTTCGGGAATTGGACCACTTTCAGACGTTGGTCAATGATGCTCAAAAACGGATCA 466
Db 7202777 ACTTCGGTACCAACAGGACCTCGAAACCCCTCATCTCCAGGGCGCACGCGCAGGGCATGA 7202836

Qy 467 AGTGATTGTGCACTTTGTGCCCAATCA 494
Db 7202837 AGGTCTTCTGACGTCATCACCAACCA 7202864

RESULT 7
US-10-213-990-41
; Sequence 41, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUNIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Aspergillus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1488)
US-10-213-990-41

Query Match 2.3%; Score 49; DB 14; Length 1488;
Best Local Similarity 48.9%; Pred. No. 0.00012;
Matches 161; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy 234 GTATTGGGGGGGATCTGGAGGGGTTCTCAAAAATTCCTTATCTTAAACAGCTGGG 293
Db 165 GTATTGGTGTGTACATGCGAAGGCATTATTGAACAACTTCGATTACATCAAGGCGATGG 224

Qy 294 CGTAACGACAAATCTGTTCTCCCGGTTTGGACAATCTGGTACATCTGGCGGGCACCGA 353
Db 225 CTTTACAGCAATTGGATAACCCAG--TCACAGAGAACTTCCCGAGGATAGTCGGA 281

Qy 354 TAACACGGGCTATCAGGATACCTGACGCGCGATTTTAAACAGATTGGAGAACATTTCCG 413
Db 282 AGGCACTGCATACCAAGGCTCTGGCAGCAAGACATTTATTTCGTCATTCACCACTATGG 341

Qy 414 GAATTGGACCACTTTCAGACGTTGGTCAATGATGCTACCAAAACGGATCAAGGTAT 473
Db 342 GACCGCTGACGACCTGAAAGCCCTGGCGTCAGCTTTCATGACAGGGGCGATGTATCTCAT 401

Qy 474 TGTGCACTTTGTGCCCAATCTGCACTTCCTTTTAAAGCAACGATTCACCTTTTGGGA 533
Db 402 GGTGATGTTGTTGCCAATCAGATGGGATATGCCGTCGAGCGGATTCGTCGACTATAG 461

Qy 534 AGGCGGCGCCCTCAACAAATGGAACCT 562
Db 462 TGTCTTCAACCCCTTCAACTCCAGACCT 490

RESULT 8
US-10-228-063-46
; Sequence 46, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lanahan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12

```
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 3285
; TYPE: DNA
; ORGANISM: Aspergillus shirousami
US-10-228-063-46

Query Match      2.2%; Score 47.8; DB 12; Length 3285;
Best Local Similarity 47.8%; Pred. No. 0.00052;
Matches 139; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 940 GTGCTGATTTGATCTCAACGCGGTGATCGAATGTTGGCACATTTAACGAAACG 999
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1000 ATGTACGATCTTAACAATATGTGTGAACCAACGGGGAACGATACAAATACAAAGAAAT 1059
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 805 ATGGACGACCTCTACACATGATCAACACCGTGAAGTCGAGCTGCCGCGACTCCACCCCTC 864
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1060 CTAATCAATATTCGATACACATGATATGTCAGATTTCTTCGGTAAATTCGAACAAG 1119
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 865 CTCGCGACCTTCGTGGGAACACGACCAACCCGCGCTTCGCTCTCTACACCAACGACATC 924
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1120 GCGAATTTGCACACGCGCTTGTCTTCTCTCACTTCGCGGGGTACGCCCTCCATCTAT 1179
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 925 GCCTCGCCAGAACGTGGCGGCTTCATCTCTCAACGCGCATCCGATCATCTAC 984
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1180 TATGGAACCGAACAAATATATGCGAGCGGCAATCAACCCGTACAAACCGGGG 1230
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 985 GCGCGCCAGGACGACACTACGCGCGGCAACGACCGGCAACCGCGAG 1035
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-10-050-763-2
; Sequence 2, Application US/10050763
; Publication No. US20030135884A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Su-May
; APPLICANT: Shaw, Jui-Fu
; TITLE OF INVENTION: TRANSGENIC SEEDS EXPRESSING
; TITLE OF INVENTION: AMYLOPULLULANASE AND USES THEREOF
; FILE REFERENCE: 08919-067001
; CURRENT APPLICATION NUMBER: US/10/050,763
; CURRENT FILING DATE: 2002-01-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2863
; TYPE: DNA
; ORGANISM: Thermoana ethanolicus
US-10-050-763-2

Query Match      2.2%; Score 46.8; DB 12; Length 2863;
Best Local Similarity 57.5%; Pred. No. 0.001;
Matches 84; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 363 CTATCAGGACTTGGCGCGGATTTTAAACAGATTGAGACATTTGGGAATTGGAC 422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1124 CAATCAGGATGATTAACACCGATTACAAAGATAGACGAGTTATTGGGAGATTAGA 1183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 423 CACATTTGACACGTTGTGTAATGATGCTCACCAAAAGGAATCAAGTGATTTGTCGACTT 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1184 TACATTTAAACACTTATGAAGACCCATCGACAGGATTAAGTATATATCTGTGG 1243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 483 TGTGCCCATCATTCGCTCTTTTA 508
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1244 CGTCTTCATCATCAAGTGATGATA 1269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-09-974-300-655
; Sequence 655, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085-500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 655
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-655

Query Match      2.1%; Score 46.4; DB 10; Length 1683;
Best Local Similarity 55.6%; Pred. No. 0.00095;
Matches 89; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 222 GAAGTGGAAATGATTTGGGGCGGGGATCTGGAGGGGTTCTGCAAAACTTCCTTATCT 281
Db 60 GGATTCGAACGGGGATGGCGTCGGAGATATCAACGGGATCAGAGAGAGCTTCGGTATAT 119
QY 282 TAAACACTGGCGTAACGACAATCTGTTGTCTCCCGGTTTGGACAACTCTGGATACACT 341
Db 120 CAGGAGCTCGGGCGAGATGCCATTGGATTGTCCGGTCTTTGATTCTCCAAATGCAGA 179
QY 342 GCGCGGACCGGATACACGGGCTATCACGGATCTGACG 381
Db 180 TAACGGCTACGATCAGGATTTATCAAGCATCATCGCG 219

RESULT 11
US-10-128-714-203/c
; Sequence 203, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wendi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 203
; LENGTH: 3393
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-203

Query Match      2.1%; Score 46.2; DB 14; Length 3393;
```

RESULT 13
US-10-213-990-44
; Sequence 44, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05

```

; SEQ ID NO 44
; TYPE: DNA
; LENGTH: 1893
; ORGANISM: Aspergillus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1893)
US-10-213-990-44

Query Match                2.0%; Score 43.8; DB 14; Length 1893;
Best Local Similarity 46.6%; Ered.No.0.0074;
Matches 138; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY      968  TTCGAATGTGTTGCGGCACATTACGCAAAAGCATGTACGATCTTAAACAATATGTTGAACC 1027
          |||||
Db       842  TTCTCTACGGTTTCCAATCGACCAGCGGAGCATAGCAATCTGTACAAACATGATCAGCT 901
          |||||
QY      1028  AAACGGGACAGGTACAAATACAAAGAAATCTATACATTTATCGATPACCGTGA 1087
          |||||
Db       902  CCGTGTGGTCTGACTGTGGGATCCCACTTTGCTCGCAACTTTATCGAAGCAATGATA 961
          |||||
QY      1088  TGTCAAGATTTCTTTGGTAAATTGCAACAAGCGCAATTGCAACGAGCGCTTCTCTTCA 1147
          |||||
Db       962  ACCCCGATTTGCTCTCTTATACGAGCGACTATTTCGCAAGCAAGAACTCATCTCTTCA 1021
          |||||
QY      1148  TTCTCACTTCGGGGTAGCGCTTCCATCTTATTATGGRACCGAACAATACATGCGAGGCG 1207
          |||||
Db       1022  TGTTTCTTCCAGGCTATCCCAATGTCTACGCCCGAGGAGCAGCACTACAGCGGCG 1081
          |||||
QY      1208  GCAATGACCGGTACAAACCGGGGGATGATCGCGCGTTTGATACGACAACCAACCGC 1262
          |||||
Db       1082  GTGTGACCTGTCCAAACCGGAGGCTGTCTGGTGTCTGGATCTCTGACACTCGACAGGCG 1136
          |||||

```

```

00      GIGCIGACCCGCGCACACAGGCCTGTCTGGCGTGTCTGATCTGACCAAGCGC 1136

RESULT 14
US-08-781-986A-56
; Sequence 56, Application US/08781.986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: Hp Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.986A
; FILING DATE:
```



```
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 30246 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-56

Query Match      2.0%; Score 43.2; DB 8; Length 30246;
Best Local Similarity 57.4%; Pred. No. 0.078;
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 365 ATACGGATACCTGGACCGCGATTTTAAACAGATTGAGGAACATTTTCGGGAATTGGACCA 424
DB 12466 ATATGGCTATGATATCAGCAATTATTAGAAATCAATGAAGACTTTTGGACGATGGATG 12525

QY 425 CATTTGACACCTTGGTCAATGATGCTCCACCAAGGAATCAAGGTGATTGTCGACTTTG 484
DB 12526 ATTTGAAAGCTTAATCAAGTTGGCATCAAAAAGACTTGAAGATGATGTTAGATATTG 12585

QY 485 TGCCCAATCATTCGAC 500
DB 12586 TCATTATCATACGTC 12601

RESULT 15
US-10-081-872-185
Sequence 185, Application US/10081872
Publication No. US20030125534A1
GENERAL INFORMATION:
APPLICANT: Callen, Walter
APPLICANT: Richardson, Toby
APPLICANT: Frey, Gerhard
APPLICANT: Short, Jay M.
APPLICANT: Mathur, Eric J.
APPLICANT: Gray, Kevin A.
APPLICANT: Kerovuo, Janne S.
APPLICANT: Slupska, Malgorzata
TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
TITLE OF INVENTION: AND METHODS OF USE THEREOF
FILE REFERENCE: 09010-108001
CURRENT APPLICATION NUMBER: US/10/081,872
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: US 60/270,495
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: US 60/270,496
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: US 60/291,122
PRIOR FILING DATE: 2001-05-14
NUMBER OF SEQ ID NOS: 321
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 185
LENGTH: 1572
TYPE: DNA
ORGANISM: Environmental
US-10-081-872-185

Query Match      2.0%; Score 42.2; DB 14; Length 1572;
Best Local Similarity 54.1%; Pred. No. 0.022;
Matches 86; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 341 TGCGGGCACCGGATACACGGGCTATCACGGATACGCGCGATTTTAAACAGATTG 400
```

```
DB 278 TGCCAGTCACCCCTTCTCTAGCTATCATAAATATGATCTAACGGAATTATATACATTG 337
QY 401 AGGAACATTTTCGGGAATTGGACCAACATTTGACACGTTGGTCAATGATGCTCACCAAAAG 460
DB 338 ATCCGCACTACGGAATCTGCAAGATTTTCGCAAGCTGATGAAGAGACAGACAAACGAG 397
QY 461 GAATCAAGGTGATTGTCGACTTTTGCCCAATCAATTCGA 499
DB 398 ACGTAAAGTCATTATGGACCTTTGTGTGAATCATACGA 436
```

Search completed: September 17, 2003, 12:17:53
Job time : 395 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 08:16:17 ; Search time 3058 Seconds
(without alignments)
17167.321 Million cell updates/sec

Title: US-10-069-908-1
Perfect score: 2160
Sequence: 1 atgaaaagaaacgcttcc.....ctgttacgtggcaaaatag 2160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	49	2.3	362 9	AA787426
2	48.8	2.3	464 9	AA965792
3	48.8	2.3	541 9	AA787396
4	48.8	2.3	583 9	AA786071

5	46	2.1	522	12	BM646781	BM646781	170006873
6	46	2.1	547	12	BM636880	BM636880	170006875
7	46	2.1	554	12	BM630193	BM630193	170006875
8	46	2.1	612	11	CN9092VR	CN9092VR	170006875
9	46	2.1	617	12	BM647914	BM647914	170006875
10	46	2.1	632	12	BM583892	BM583892	170006872
11	46	2.1	634	12	BM603238	BM603238	170006870
12	46	2.1	654	12	BM578896	BM578896	170006873
13	46	2.1	657	12	BM588117	BM588117	170006873
14	46	2.1	664	12	BM629008	BM629008	170006874
15	46	2.1	665	12	BM624708	BM624708	170006874
16	46	2.1	665	12	BM652834	BM652834	170006873
17	46	2.1	679	12	BM556725	BM556725	170006873
18	46	2.1	681	12	BM614246	BM614246	170006871
19	46	2.1	684	12	BM620803	BM620803	170006874
20	46	2.1	684	12	BM634536	BM634536	170006875
21	46	2.1	684	12	BM635945	BM635945	170006875
22	46	2.1	694	12	BM579790	BM579790	170006872
23	46	2.1	698	12	BM616688	BM616688	170006871
24	46	2.1	704	12	BM628529	BM628529	170006874
25	46	2.1	706	12	BM650445	BM650445	170006873
26	45.8	2.1	655	14	CD033125	CD033125	mgmt-011xL
27	44.2	2.0	560	12	BJ413818	BJ413818	BJ410968
28	44.2	2.0	601	12	BJ410968	BJ410968	BJ410968
29	44.2	2.0	636	12	BJ416571	BJ416571	BJ416571
30	44.2	2.0	668	9	AU060851	AU060851	AU060851
31	44.2	2.0	677	9	AU060198	AU060198	AU060198
32	43	2.0	867	29	CNS066F0	CNS066F0	EST 31-JUL-1998
33	42.6	2.0	1201	13	EX381961	EX381961	EST 31-JUL-1998
34	41.2	1.9	402	28	AZ877703	AZ877703	EST 31-JUL-1998
35	41	1.9	477	13	BQ102620	BQ102620	6HRm161.6
36	40.8	1.9	841	29	CNS010LW	CNS010LW	AL099134 Drosophil
37	40.8	1.9	885	13	BX425603	BX425603	AL099134 Drosophil
38	40.6	1.9	304	9	AI210530	AI210530	BX425603
39	40.2	1.9	553	9	AU262360	AU262360	AI210530 i7h08al.r
40	40	1.9	529	9	AI406075	AI406075	AU262360
41	40	1.9	530	9	AA539984	AA539984	AI406075 GH26261.5
42	40	1.9	548	9	AI386686	AI386686	AA539984 LD19296.5
43	40	1.9	574	9	AI402499	AI402499	AI386686 GH16849.5
44	40	1.9	577	9	AI134329	AI134329	AI402499 GH21734.5
45	40	1.9	596	12	BI373231	BI373231	AI134329 GH11763.5

ALIGNMENTS

RESULT 1
AA787426
LOCUS
DEFINITION
AA787426 362 bp mRNA linear EST 31-JUL-1998
n3c10a1.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone n3c10a1, mRNA sequence.
ACCESSION
AA787426
VERSION
AA787426.1 GI:2847657
KEYWORDS
EST.
SOURCE
Emericella nidulans (anamorph: Aspergillus nidulans)
ORGANISM
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emericella.
REFERENCE
1 (bases 1 to 362)
AUTHORS
Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R., Prade, R. and Roe, B.
TITLE
An Aspergillus nidulans EST Database
JOURNAL
Unpublished
COMMENT
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center

```

FEATURES
  source
    High quality sequence stop: 346.
    Location/Qualifiers
      1..362
        /organism="Emericella nidulans"
        /mol_type="mRNA"
        /strain="FGSC A26"
        /db_xref="taxon:162425"
        /clone="n3c10a1"
        /tissue_type="vegetative mycelia, asexual structures"
        /clone_lib="Aspergillus nidulans 24hr asexual
        developmental and vegetative cDNA lambda zap library"
        /note="vector: pBluescript SK-; Site 1: EcoRI; Site 2:
        XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
        3' end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT      92 a 117 c 86 g 67 t
ORIGIN
  Query Match      2.3%; Score 49; DB 9; Length 362;
  Best Local Similarity 51.1%; Pred. No. 0.0031;
  Matches 115; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 1004 ACATCTTAAACAATATGTTGAACAAACGGGGAACGAGTACAAATACAAGAAATCTAA 1063
DB 51 AGATATCTCGCATGATGACAGTGCATGATGCTCTCCCGATCCAGGCCCTCG 110
QY 1064 TCATTTATCATTAACATGATATGCAAGATTTCTTCGGTAAATTCGAACAGGGCA 1123
DB 111 GCATTCATGACAAACACGACATGCACGCTGGCTGATTAATAGGACGCAAGTCTC 170
QY 1124 ATTTCACACGCGCTTGCTTTCATCTCATCTCCGGGGTACGCCCTCCATCTATTATG 1183
DB 171 TCTGGAAGAACCACTGCCATGATGATTCGACAGGGATTCCTCATGCTACTACG 230
QY 1184 GAACCGAACAATPACATGGCAGGCGCAATGACCCGTACAAACCGG 1228
DB 231 GGACAGACAGCGGTACCGGGTGGCAATGATCCGAGAACCGCG 275

RESULT 2
AA965792
LOCUS
  DEFINITION
    o5e03a1.r1 Aspergillus nidulans 24hr asexual developmental and
    vegetative cDNA lambda zap library Emericella nidulans cDNA clone
    o5e03a1.5', mRNA sequence.
  ACCESSION
    AA965792
  VERSION
    AA965792.1 GI:3139676
  KEYWORDS
    EST.
  ORGANISM
    Emericella nidulans (anamorph: Aspergillus nidulans)
    Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
    Eurotiales; Trichocomaceae; Emericella.
  REFERENCE
    1 (bases 1 to 464)
    Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
    Prade, R. and Roe, B.
  TITLE
    An Aspergillus nidulans EST Database
  JOURNAL
    Unpublished
  COMMENT
    Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
    Department of Chemistry and Biochemistry
    Advanced Center for Genome Technology, University of Oklahoma
    620 Parrington Oval, Norman, OK 73019, USA
    Tel: 405 325 4912
    Fax: 405 325 7762
    Email: broe@ou.edu
    We anticipate the future release of the cDNA clones to the Fungal
    Genetics Stock Center
    Seq primer: T3
    High quality sequence stop: 453.
    Location/Qualifiers
      1..464
        /organism="Emericella nidulans"
        /mol_type="mRNA"
        /strain="FGSC A26"
        /db_xref="taxon:162425"

```

```

/clone="o5e03a1"
/tissue_type="vegetative mycelia, asexual structures"
/clone_lib="Aspergillus nidulans 24hr asexual
developmental and vegetative cDNA lambda zap library"
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
3' end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT      118 a 135 c 94 g 117 t
ORIGIN
  Query Match      2.3%; Score 48.8; DB 9; Length 464;
  Best Local Similarity 48.8%; Pred. No. 0.004;
  Matches 134; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 968 TTCGAATATGTTTCGGCACATTTAGCCAAACGATGATGATCTTAACAATATGTTGAACC 1027
DB 67 TTCTGAATGCGTTCAAGTCTCTCGAGCGGAGCATGTCGGATCTCTATAACATGATCAACA 126
QY 1028 AAAACGGGACGAGTACAAATACAAGAAATCTAATACATTTATGATTAACCATGATA 1087
DB 127 CAGTCGCTCAAAATTCGCGGATCCTACATGCTTGGAACTTTATCGAGAACCATGACA 186
QY 1088 TGTCAAGATTTCTTCGGTAAATTCGAACAAGCGCAATTTGCACCGCGCGTTGCTTTCA 1147
DB 187 ATCTCTCGATTTCCCAACTATCTCCGGTATGATCGGCGCAAGACGTCTCTCGGTTCC 246
QY 1148 TTCTCATCTTCGCGGATGAGCCCTCCATCTATTATGGAACCGCAACATACATGCGAGGCG 1207
DB 247 TCTTCTTGAACCGGGAATPCCCTATTGTTTATCGGCGCAGGAGCAACACTATTTCAGGCA 306
QY 1208 GCATGATACCGGTACAAACGGGCGGATGATCGCGCGT 1243
DB 307 GCATGATACCGGTATTAACCGGCGCGGTTGGTGGT 342

RESULT 3
AA787396
LOCUS
  DEFINITION
    n3a12a1.r1 Aspergillus nidulans 24hr asexual developmental and
    vegetative cDNA lambda zap library Emericella nidulans cDNA clone
    n3a12a1, mRNA sequence.
  ACCESSION
    AA787396
  VERSION
    AA787396.1 GI:2847627
  KEYWORDS
    EST.
  ORGANISM
    Emericella nidulans (anamorph: Aspergillus nidulans)
    Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
    Eurotiales; Trichocomaceae; Emericella.
  REFERENCE
    1 (bases 1 to 541)
    Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
    Prade, R. and Roe, B.
  TITLE
    An Aspergillus nidulans EST Database
  JOURNAL
    Unpublished
  COMMENT
    Other ESTs: n3a12a1.f1
    Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
    Department of Chemistry and Biochemistry
    Advanced Center for Genome Technology, University of Oklahoma
    620 Parrington Oval, Norman, OK 73019, USA
    Tel: 405 325 4912
    Fax: 405 325 7762
    Email: broe@ou.edu
    We anticipate the future release of the cDNA clones to the Fungal
    Genetics Stock Center
    High quality sequence stop: 510.
    Location/Qualifiers
      1..541
        /organism="Emericella nidulans"
        /mol_type="mRNA"
        /strain="FGSC A26"
        /db_xref="taxon:162425"
        /clone="n3a12a1"
        /tissue_type="vegetative mycelia, asexual structures"
        /clone_lib="Aspergillus nidulans 24hr asexual

```

developmental and vegetative cDNA lambda zap library"
/note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
3' end of cDNA cloned into XhoI site of pBluescript"

BASE COUNT 135 a 153 c 116 g 137 t
ORIGIN
Query Match 2.3%; Score 48.8; DB 9; Length 541;
Best Local Similarity 48.6%; Pred. No. 0.0042;
Matches 134; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 968 TTCGAATGCTTCGGCACATTTAGCAACGAGTACGATCTTAAACATATGGTGAACC 1027
DB 122 TTCTGAATGCTTCGAATGCTTCGGCACGAGCATGTGGATCTCTATACATGATCAACA 181
QY 1028 AAACGGGAACGAGTACAAATACAAAGAAATCTAAATACATTTATCGATACCATGATA 1087
DB 182 CAGTCGCTCAAAATGTCGGGATCTTACACTGCTTGGAAACTTTATCGAAGACCATGACA 241
QY 1088 TGTCAGATTTCTTTTCGGTAAATTCGAACAGGCGAATTTGCACCGGCGTTCCTTTCA 1147
DB 242 ATCTCTGATTTCCCACTATCTCCGGATATGAGTCGGGCCAAGAACGTCCTCGCGTTCC 301
QY 1148 TTCTCACTTCGGGGGTAGCCCTCCATCTATTATGGAACCGAACAATACATGCGAGCG 1207
DB 302 TCTTCTTGAACGAGCATCTCTTATTTGTTATGCGGCCAGAGCACTATTCAGGCA 361
QY 1208 GCAATGACCGGTACACCGGGGATGATCCGGCGT 1243
DB 362 GCAATGATCCCTATACCGGGAGCGCGTTTGGTGGT 397

RESULT 4
AA786071 583 bp mRNA linear EST 31-JUL-1998
LOCUS
DEFINITION J4h07a1.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emicella nidulans cDNA clone J4h07a1 5', mRNA sequence.
ACCESSION AA786071
VERSION
KEYWORDS
SOURCE Emicella nidulans (anamorph: Aspergillus nidulans)
ORGANISM Emicella nidulans
REFERENCE Eukaryota; Fungi; Ascomycota; Peizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emicella.
AUTHORS Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,B.
TITLE An Aspergillus nidulans EST Database
JOURNAL Unpublished
COMMENT Other ESTs: j4h07a1.f1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
Seq primer: SK
High quality sequence stop: 526.
FEATURES
Location/Qualifiers
1..583
/organism="Emicella nidulans"
/mol_type="mRNA"
/strain="FGSC A26"
/db_xref="taxon:162425"
/clone="j4h07a1"
/tissue_type="vegetative mycelia, asexual structures"
/clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
/note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
3' end of cDNA cloned into XhoI site of pBluescript"

BASE COUNT 143 a 168 c 124 g 148 t
ORIGIN
Query Match 2.3%; Score 48.8; DB 9; Length 583;
Best Local Similarity 48.6%; Pred. No. 0.0044;
Matches 134; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 968 TTCGAATGCTTCGGCACATTTAGCAACGAGTACGATCTTAAACATATGGTGAACC 1027
DB 120 TTCTGAATGCTTCGAATGCTTCGGCACGAGCATGTGGATCTCTATACATGATCAACA 179
QY 1028 AAACGGGAACGAGTACAAATACAAAGAAATCTAAATACATTTATCGATACCATGATA 1087
DB 180 CAGTCGCTCAAAATGTCGGGATCTTACACTGCTTGGAAACTTTATCGAAGACCATGACA 239
QY 1088 TGTCAGATTTCTTTTCGGTAAATTCGAACAGGCGAATTTGCACCGGCGTTCCTTTCA 1147
DB 240 ATCTCTGATTTCCCACTATCTCCGGATATGAGTCGGGCCAAGAACGTCCTCGCGTTCC 299
QY 1148 TTCTCACTTCGGGGGTAGCCCTCCATCTATTATGGAACCGAACAATACATGCGAGCG 1207
DB 300 TCTTCTTGAACGAGCATCTCTTATTTGTTATGCGGCCAGAGCACTATTCAGGCA 359
QY 1208 GCAATGACCGGTACACCGGGGATGATCCGGCGT 1243
DB 360 GCAATGATCCCTATACCGGGAGCGCGTTTGGTGGT 395

RESULT 5
BM646781 522 bp mRNA linear EST 26-FEB-2002
LOCUS
DEFINITION 1700687322896 A.Gam.ad.cDNA1 Anopheles gambiae cDNA clone 1960049619250 5', mRNA sequence.
ACCESSION BM646781
VERSION
KEYWORDS BM646781.1 GI:18946292
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
REFERENCE 1..522
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celerera Anopheles gambiae EST project
JOURNAL Unpublished
COMMENT Contact: Holt R.A.
Celerera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltR@celera.com
Plate: N701004NAL row: E column: 08
Seq primer: M13 Reverse.

FEATURES
Location/Qualifiers
1..522
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)"
/db_xref="taxon:7165"
/clone="1960049619250"
/dev_stage="Adult"
/lab_host="DHI0b"
/clone_lib="A.Gam.ad.cDNA1"
/note="Vector: pSport1; Site_1: SalI; Site_2: NotI; Whole adult mosquitoes (mixed sex) frozen on liquid nitrogen. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)."

```

BASE COUNT      113 a      117 c      160 g      132 t
ORIGIN
Query Match
Best Local Similarity 2.1%; Score 46; DB 12; Length 522;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 358 ACGGGCTATCAGGATCTGACGCGGATTTTAAACAGATTGAGGAACATTTCGGGAAT 417
Db 276 ATGGTCGATTTGGCTATGATTTCTGACTTTCGGACATTACAGGAGTTTGGACG 335

Qy 418 TGGACCACATTGACAGCTTGTCATATGCTCACCACAAACGGAATCAAGGTGATTGTC 477
Db 336 ATGGCTGACTTTGAGCGGTTGGTGAGCAGGCGCAGCGGCTCGGCCTGAGGTGATCATG 395

Qy 478 GACTTTGTGCCCATCATTCGA 499
Db 396 GACTTTGTGCCGAACCATTCGA 417

RESULT 6
BM636880
LOCUS
DEFINITION
17000687563021 A.Gam.ad.cdNA1 Anopheles gambiae cdNA clone
19600449641917 5', mRNA sequence.
ACCESSION
BM636880
VERSION
BM636880.1 GI:18936391
KEYWORDS
EST.
SOURCE
Anopheles gambiae (African malaria mosquito)
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE
1 (bases 1 to 547)
Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
Celera Genomics
Celera Genomics
Contact: Holt R.A.
45 w. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltR@celera.com
Plate: NU01003CY2 row: E column: 19
Seq primer: M13 Reverse.
LOCATION/Qualifiers
1..547
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449641917"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A.Gam.ad.cdNA1"
/notes="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen.
cDNA inserts >500 bp cloned directionally into pSport 1.
Not 1 site is 3'. Clones available through the Malaria
Research and Reference Reagent Resource Center
(www.malaria.mr4.org)."
BASE COUNT      123 a      122 c      165 g      137 t
ORIGIN
Query Match
Best Local Similarity 2.1%; Score 46; DB 12; Length 547;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 358 ACGGGCTATCAGGATCTGACGCGGATTTTAAACAGATTGAGGAACATTTCGGGAAT 417
Db 280 ATGGTCGATTTGGCTATGATTTCTGACTTTCGGACATTACAGGAGTTTGGACG 339

BASE COUNT      113 a      117 c      160 g      132 t
ORIGIN
Query Match
Best Local Similarity 2.1%; Score 46; DB 12; Length 522;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 358 ACGGGCTATCAGGATCTGACGCGGATTTTAAACAGATTGAGGAACATTTCGGGAAT 417
Db 276 ATGGTCGATTTGGCTATGATTTCTGACTTTCGGACATTACAGGAGTTTGGACG 335

Qy 418 TGGACCACATTGACAGCTTGTCATATGCTCACCACAAACGGAATCAAGGTGATTGTC 477
Db 336 ATGGCTGACTTTGAGCGGTTGGTGAGCAGGCGCAGCGGCTCGGCCTGAGGTGATCATG 395

Qy 478 GACTTTGTGCCCATCATTCGA 499
Db 396 GACTTTGTGCCGAACCATTCGA 417

RESULT 7
BM630193
LOCUS
DEFINITION
17000687501027 A.Gam.ad.cdNA1 Anopheles gambiae cdNA clone
19600449638734 5', mRNA sequence.
ACCESSION
BM630193
VERSION
BM630193.1 GI:18929704
KEYWORDS
EST.
SOURCE
Anopheles gambiae (African malaria mosquito)
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE
1 (bases 1 to 554)
Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
Celera Genomics
Celera Genomics
Contact: Holt R.A.
45 w. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltR@celera.com
Plate: NU01004AYR row: A column: 04
Seq primer: M13 Reverse.
LOCATION/Qualifiers
1..554
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449638734"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A.Gam.ad.cdNA1"
/notes="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen.
cDNA inserts >500 bp cloned directionally into pSport 1.
Not 1 site is 3'. Clones available through the Malaria
Research and Reference Reagent Resource Center
(www.malaria.mr4.org)."
BASE COUNT      123 a      126 c      165 g      140 t
ORIGIN
Query Match
Best Local Similarity 2.1%; Score 46; DB 12; Length 554;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 358 ACGGGCTATCAGGATCTGACGCGGATTTTAAACAGATTGAGGAACATTTCGGGAAT 417
Db 276 ATGGTCGATTTGGCTATGATTTCTGACTTTCGGACATTACAGGAGTTTGGACG 335

Qy 418 TGGACCACATTGACAGCTTGTCATATGCTCACCACAAACGGAATCAAGGTGATTGTC 477
Db 336 ATGGCTGACTTTGAGCGGTTGGTGAGCAGGCGCAGCGGCTCGGCCTGAGGTGATCATG 395

Qy 478 GACTTTGTGCCCATCATTCGA 499
Db 396 GACTTTGTGCCGAACCATTCGA 417

```

```

RESULT 8
CNS092YR      612 bp  mRNA  linear  HTC 07-JAN-2003
LOCUS         Single read from an extremity of a full-length cDNA clone made from
DEFINITION    Anopheles gambiae total adult females. 5-PRIME end of clone
               FK0A010BG11 of strain 6-9 of Anopheles gambiae (African malaria
               mosquito).
ACCESSION     BX039919
VERSION       BX039919.1  GI:27613200
KEYWORDS      HTC.
SOURCE        Anopheles gambiae (African malaria mosquito)
ORGANISM      Anopheles gambiae
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
               Anopheles.
REFERENCE     1 (bases 1 to 612)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage :
               BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
               - Web : www.genoscope.cns.fr)
FEATURES     Location/Qualifiers
               1..612
               /organism="Anopheles gambiae"
               /mol_type="mRNA"
               /strain="6-9"
               /db_xref="taxon:7165"
               /clone="FK0A010BG11"
               /plasmid="pME18-FL"
               /notes="end : 5-PRIME"
BASE COUNT   142 a 133 c 183 g 154 t
ORIGIN
               1..612
               /organism="Anopheles gambiae"
               /mol_type="mRNA"
               /strain="6-9"
               /db_xref="taxon:7165"
               /clone="FK0A010BG11"
               /plasmid="pME18-FL"
               /notes="end : 5-PRIME"
Query Match      2.1%; Score 46; DB 11; Length 612;
Best Local Similarity 57.7%; Pred. No. 0.032;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 358 ACGGCTATACCGGATCTGGACGGCGGATTTAAACAGATTCAGGACATTCGGGAAT 417
      |||||
Db 278 ATGGTCGATTTGGCTATGATATTTCTGACITTCGGGACATTCACGAGGATTTGGACG 337

QY 418 TGGACACATTTGACAGTTGGTCATGATGCTCACCACAAACGGAATCAAGTGATGTC 477
      |||||
Db 338 ATGGCTGACTTTGAGCGTTGGTGGACGAGCGCGGCTCGGCTGAAGTGATCATG 397

QY 478 GACTTTGGCCCAATCATTCGA 499
      |||||
Db 398 GACTTTGGCCGACCATTCGA 419

RESULT 9
BM647914      617 bp  mRNA  linear  EST 26-FEB-2002
LOCUS         17000687324532 A.Gam.ad.cdna1 Anopheles gambiae cDNA clone
DEFINITION    19600449658038 5', mRNA sequence.
ACCESSION     BM647914
VERSION       BM647914.1  GI:18947425
KEYWORDS      EST.
SOURCE        Anopheles gambiae (African malaria mosquito)
ORGANISM      Anopheles gambiae
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
               Anopheles.
REFERENCE     1 (bases 1 to 617)
AUTHORS      Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab
               R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
TITLE        Celera Anopheles gambiae EST project
JOURNAL      Unpublished
COMMENT       Contact: Holt R.A.
               Celera Genomics
               45 W. Gude Dr., Rockville, MD 20850, USA
               Tel: 2404533151
               Fax: 2404534580

BM647914      617 bp  mRNA  linear  EST 26-FEB-2002
LOCUS         17000687324532 A.Gam.ad.cdna1 Anopheles gambiae cDNA clone
DEFINITION    19600449658038 5', mRNA sequence.
ACCESSION     BM647914
VERSION       BM647914.1  GI:18947425
KEYWORDS      EST.
SOURCE        Anopheles gambiae (African malaria mosquito)
ORGANISM      Anopheles gambiae
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
               Anopheles.
REFERENCE     1 (bases 1 to 617)
AUTHORS      Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab
               R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
TITLE        Celera Anopheles gambiae EST project
JOURNAL      Unpublished
COMMENT       Contact: Holt R.A.
               Celera Genomics
               45 W. Gude Dr., Rockville, MD 20850, USA
               Tel: 2404533151
               Fax: 2404534580

```

```

Email: HoltRA@celera.com
Plate: NU01004HNK row: E column: 12
Seq primer: M13 Reverse.
Location/Qualifiers
               1..617
               /organism="Anopheles gambiae"
               /mol_type="mRNA"
               /strain="RSP-ST (Reduced susc. to Permethrin - std.
               chromosome)"
               /db_xref="taxon:7165"
               /clone="19600449658038"
               /dev_stage="Adult"
               /lab_host="DH10b"
               /clone_lib="A.Gam.ad.cdna1"
               /note="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole
               adult mosquitoes (mixed sex) frozen on liquid nitrogen.
               cDNA inserts >500 bp cloned directionally into pSport 1.
               Not 1 site is 3'. Clones available through the Malaria
               Research and Reference Reagent Resource Center
               (www.malaria.wz4.org)."
BASE COUNT   139 a 145 c 183 g 150 t
ORIGIN
               1..617
               /organism="Anopheles gambiae"
               /mol_type="mRNA"
               /strain="RSP-ST (Reduced susc. to Permethrin - std.
               chromosome)"
               /db_xref="taxon:7165"
Query Match      2.1%; Score 46; DB 12; Length 617;
Best Local Similarity 57.7%; Pred. No. 0.032;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 358 ACGGCTATACCGGATCTGGACGGCGGATTTAAACAGATTCAGGACATTCGGGAAT 417
      |||||
Db 267 ATGGTCGATTTGGCTATGATATTTCTGACITTCGGGACATTCACGAGGATTTGGACG 326

QY 418 TGGACACATTTGACAGTTGGTCATGATGCTCACCACAAACGGAATCAAGTGATGTC 477
      |||||
Db 327 ATGGCTGACTTTGAGCGTTGGTGGACGAGCGCGGCTCGGCTGAAGTGATCATG 386

QY 478 GACTTTGGCCCAATCATTCGA 499
      |||||
Db 387 GACTTTGGCCGACCATTCGA 408

RESULT 10
BM583892      632 bp  mRNA  linear  EST 22-FEB-2002
LOCUS         17000687279414 A.Gam.ad.cdna.blood1 Anopheles gambiae cDNA clone
DEFINITION    19600449707142 5', mRNA sequence.
ACCESSION     BM583892
VERSION       BM583892.1  GI:18872359
KEYWORDS      EST.
SOURCE        Anopheles gambiae (African malaria mosquito)
ORGANISM      Anopheles gambiae
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
               Anopheles.
REFERENCE     1 (bases 1 to 632)
AUTHORS      Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab
               R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
TITLE        Celera Anopheles gambiae EST project
JOURNAL      Unpublished
COMMENT       Contract: Holt R.A.
               Celera Genomics
               45 W. Gude Dr., Rockville, MD 20850, USA
               Tel: 2404533151
               Fax: 2404534580
               Email: HoltRA@celera.com
               Place: NU010049V2 row: B column: 24
               Seq primer: M13 Reverse.
               Location/Qualifiers
               1..632
               /organism="Anopheles gambiae"
               /mol_type="mRNA"
               /strain="RSP-ST (Reduced susc. to Permethrin - std.
               chromosome)"
               /db_xref="taxon:7165"
FEATURES     Location/Qualifiers
               1..632
               /organism="Anopheles gambiae"
               /mol_type="mRNA"
               /strain="RSP-ST (Reduced susc. to Permethrin - std.
               chromosome)"
               /db_xref="taxon:7165"

```

```

/clone="19600449707142"
/dev_stage="Adult"
/lab_host="DH10b"
/notes="Vector: pSport1; Site 1: Sall; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"
BASE COUNT      140 a      152 c      189 g      151 t
ORIGIN

Query Match      2.1%; Score 46; DB 12; Length 632;
Best Local Similarity 57.7%; Pred. No. 0.033;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 358 ACGGCTATACCGATCTGAGCGCGGATTTTAAACAGATGAGGAACATTTCGGGAAT 417
DB 281 ATGGTCGATTTGGCTATGATATTCTGACTTTCGGGACATTCACGAGGAGTTTGGAAAG 340
QY 418 TGGACACATTTGACACGTTGGTCAATGATGCTCACCACAAACGGAATCAAGTGATTC 477
DB 341 ATGGCTGACTTTGAGCGGTTGGTGAGCGGCGACGCGCTCGGCTGAGGTGATCATG 400
QY 478 GACTTTGGCCCAATCATTCGA 499
DB 401 GACTTTGGCCCAACCATTCGA 422

RESULT 11
BM603238
LOCUS      634 bp      mRNA      linear      EST 25-FEB-2002
DEFINITION 17000687072486 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone
BM603238
ACCESSION 19600449682521 5', mRNA sequence.
VERSION
KEYWORDS
SOURCE
ORGANISM  Anopheles gambiae (African malaria mosquito)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1 (bases 1 to 634)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
Celera Anopheles gambiae EST project
Unpublished
Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU010049V8 row: A column: 03
Seq primer: M13 Reverse.
Location/Qualifiers
1..634
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449682521"
/dev_stage="Adult"
/lab_host="DH10b"
/notes="Vector: pSport1; Site 1: Sall; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"
BASE COUNT      140 a      152 c      189 g      151 t
ORIGIN

Query Match      2.1%; Score 46; DB 12; Length 632;
Best Local Similarity 57.7%; Pred. No. 0.033;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 358 ACGGCTATACCGATCTGAGCGCGGATTTTAAACAGATGAGGAACATTTCGGGAAT 417
DB 281 ATGGTCGATTTGGCTATGATATTCTGACTTTCGGGACATTCACGAGGAGTTTGGAAAG 340
QY 418 TGGACACATTTGACACGTTGGTCAATGATGCTCACCACAAACGGAATCAAGTGATTC 477
DB 341 ATGGCTGACTTTGAGCGGTTGGTGAGCGGCGACGCGCTCGGCTGAGGTGATCATG 400
QY 478 GACTTTGGCCCAATCATTCGA 499
DB 401 GACTTTGGCCCAACCATTCGA 422

RESULT 11
BM603238
LOCUS      634 bp      mRNA      linear      EST 25-FEB-2002
DEFINITION 17000687072486 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone
BM603238
ACCESSION 19600449682521 5', mRNA sequence.
VERSION
KEYWORDS
SOURCE
ORGANISM  Anopheles gambiae (African malaria mosquito)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1 (bases 1 to 634)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
Celera Anopheles gambiae EST project
Unpublished
Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU010049V8 row: A column: 03
Seq primer: M13 Reverse.
Location/Qualifiers
1..634
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449682521"
/dev_stage="Adult"
/lab_host="DH10b"
/notes="Vector: pSport1; Site 1: Sall; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"
BASE COUNT      140 a      152 c      189 g      151 t
ORIGIN

Query Match      2.1%; Score 46; DB 12; Length 632;
Best Local Similarity 57.7%; Pred. No. 0.033;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 358 ACGGCTATACCGATCTGAGCGCGGATTTTAAACAGATGAGGAACATTTCGGGAAT 417
DB 281 ATGGTCGATTTGGCTATGATATTCTGACTTTCGGGACATTCACGAGGAGTTTGGAAAG 340
QY 418 TGGACACATTTGACACGTTGGTCAATGATGCTCACCACAAACGGAATCAAGTGATTC 477
DB 341 ATGGCTGACTTTGAGCGGTTGGTGAGCGGCGACGCGCTCGGCTGAGGTGATCATG 400
QY 478 GACTTTGGCCCAATCATTCGA 499
DB 401 GACTTTGGCCCAACCATTCGA 422

RESULT 12
BM578896
LOCUS      654 bp      mRNA      linear      EST 22-FEB-2002
DEFINITION 17000687166619 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone
BM578896
ACCESSION 19600449680544 5', mRNA sequence.
VERSION
KEYWORDS
SOURCE
ORGANISM  Anopheles gambiae (African malaria mosquito)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1 (bases 1 to 654)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
Celera Anopheles gambiae EST project
Unpublished
Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU010049V8 row: N column: 18
Seq primer: M13 Reverse.
Location/Qualifiers
1..654
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449680544"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A.Gam.ad.cDNA.blood1"
/notes="Vector: pSport1; Site 1: Sall; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"
BASE COUNT      149 a      155 c      192 g      158 t
ORIGIN

Query Match      2.1%; Score 46; DB 12; Length 654;
Best Local Similarity 57.7%; Pred. No. 0.033;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 358 ACGGCTATACCGATCTGAGCGCGGATTTTAAACAGATGAGGAACATTTCGGGAAT 417
DB 273 ATGGTCGATTTGGCTATGATATTCTGACTTTCGGGACATTCACGAGGAGTTTGGAAAG 332

```

```

QY 418 TGGACCACTTTGACAGCTTGTGCTCAATGATGCTCACCACAAACGGAATCAAGTGATTGTC 477
D5 333 ATGGCTGACCTTTGAGCGGTGTGTGGAGCAGCGCACGGCTCGCCCTGAAGGTGATCATG 392
QY 478 GACTTTGTGCCAATCATTCGA 499
D5 393 GACTTTGTGCCAATCATTCGA 414

RESULT 13
BM581117
LOCUS
DEFINITION 657 bp mRNA linear EST 25-FEB-2002
17000687322054 A.Gam.ad.cdNAL bloodl Anopheles gambiae cDNA clone
19600449685008 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Anopheles gambiae (African malaria mosquito)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE
1 (bases 1 to 657)
Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab
R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
Celera Anopheles gambiae EST project
Unpublished
Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004B2U row: H column: 18
Seq primer: M13 Reverse.
JOURNAL
TITLE
COMMENT
FEATURES
source
1..657
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="PSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449685008"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A.Gam.ad.cdNAL.bloodl"
/note="Vector: pSport1; Site 1: Sall; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria. Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"
BASE COUNT 151 a 155 c 192 g 159 t
ORIGIN
Query Match 2.1%; Score 46; DB 12; Length 657;
Best Local Similarity 57.7%; Pred. No. 0.033;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 358 ACGGCTATCAGGATCTGAGCGCGGATTTTAAACAGATTGAGGAACATTTCCGGAAAT 417
D5 278 ATGTCGATTTGGCTATGATATTTCTGACCTTCGGGACATTCACGAGGATTTGGACG 337
QY 418 TGGACCACTTTGACAGCTTGTGCTCAATGATGCTCACCACAAACGGAATCAAGTGATTGTC 477
D5 333 ATGGCTGACCTTTGAGCGGTGTGTGGAGCAGCGCACGGCTCGCCCTGAAGGTGATCATG 392
QY 478 GACTTTGTGCCAATCATTCGA 499
D5 398 GACTTTGTGCCAATCATTCGA 419

```

```

RESULT 14
BM629008
LOCUS
DEFINITION 664 bp mRNA linear EST 26-FEB-2002
17000687499420 A.Gam.ad.cdNAL Anopheles gambiae cDNA clone
19600449652279 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Anopheles gambiae (African malaria mosquito)
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE
1 (bases 1 to 664)
Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab
R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
Celera Anopheles gambiae EST project
Unpublished
Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU010049WI row: E column: 13
Seq primer: M13 Reverse.
JOURNAL
TITLE
COMMENT
FEATURES
source
1..664
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="PSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449652279"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A.Gam.ad.cdNAL"
/note="Vector: pSport1; Site 1: Sall; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen.
cDNA inserts >500 bp cloned directionally into pSport 1.
Not 1 site is 3'. Clones available through the Malaria
Research and Reference Reagent Resource Center
(www.malaria.mr4.org)."
BASE COUNT 154 a 155 c 196 g 159 t
ORIGIN
Query Match 2.1%; Score 46; DB 12; Length 664;
Best Local Similarity 57.7%; Pred. No. 0.033;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 358 ACGGCTATCAGGATCTGAGCGCGGATTTTAAACAGATTGAGGAACATTTCCGGAAAT 417
D5 281 ATGTCGATTTGGCTATGATATTTCTGACCTTCGGGACATTCACGAGGATTTGGACG 340
QY 418 TGGACCACTTTGACAGCTTGTGCTCAATGATGCTCACCACAAACGGAATCAAGTGATTGTC 477
D5 341 ATGGCTGACCTTTGAGCGGTGTGTGGAGCAGCGCACGGCTCGCCCTGAAGGTGATCATG 400
QY 478 GACTTTGTGCCAATCATTCGA 499
D5 401 GACTTTGTGCCAATCATTCGA 422
RESULT 15
BM624708
LOCUS
DEFINITION 665 bp mRNA linear EST 26-FEB-2002
17000687491707 A.Gam.ad.cdNAL Anopheles gambiae cDNA clone
19600449622270 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Anopheles gambiae (African malaria mosquito)

```


Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.

REFERENCE

1. (bases 1 to 665)
Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab
R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
Celera Anopheles gambiae EST project

TITLE

Unpublished

COMMENT

Contact: Holt R.A.

Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151

Fax: 2404534580

Email: HoltRA@celera.com

Plate: NU01004N6K row: C column: 04

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1..665
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clones="1960449622270"
/lab_host="DH10b"
/dev_stage="Adult"
/clone_lib="A.Gam.ad.cdna1"
/note="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen.
cdna inserts >500 bp cloned directionally into pSport 1.
Not I site is 3'. Clones available through the Malaria
Research and Reference Reagent Resource Center
(www.malaria.mr4.org)."

BASE COUNT 154 a 155 c 196 g 160 t

ORIGIN

Query Match 2.1%; Score 46; DB 12; Length 665;
Best Local Similarity 57.7%; Pred. No. 0.033;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 358 ACGGGTATCAGCGATCTGACGCGGATTTTAAACAGATTGAGGAACATTTCCGGGAT 417
Db |||||
283 ATGGTCGATTTGGCTATGATATTCTGACTTCGGGACATTCACGAGGAGTTGGAACG 342
QY 418 TGGACACATTTGACACGTTGGTCATGATGCTACCAAAACGGAATCAAGGTGATTC 477
Db |||||
343 ATGGCTGACTTTGAGCGGTTGGTGAGACAGCGGCTCGGCGCTGAAGGTGATCATG 402
QY 478 GACTTTGCGCCATCATTCGA 499
Db |||||
403 GACTTTGCGGACCATTCGA 424

Search completed: September 17, 2003, 10:50:06
Job time : 3062 secs